An Investigation into Passenger Car Drivers' Preferences in Loudness between Dynamic and Compressed Musical Recordings

by

Mark Stobbart

Thesis presented in partial fulfilment of the requirements for the degree of Master of Philosophy (Music Technology) in the Faculty of Music at Stellenbosch University

Supervisor: Mr. G. Roux

March 2017

Declaration

By submitting this thesis electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

Date: March 2017

Abstract

An Investigation into Passenger Car Drivers' Preferences in Loudness between Dynamic and Compressed Musical Recordings

M. Stobbart

Department of Music,
Stellenbosch University,
Private Bag X1, Matieland 7602, South Africa.

Thesis: MPhil (Music Technology)

November 2016

New international broadcasting legislation and the implementation thereof by online platforms such as YouTube and online music retailers such as iTunes, are bringing an end to the over-compressed music that has become the norm over recent years. Whilst these new loudness standards are advancing recording quality by allowing for a wider dynamic range, it may also have unintended consequences with regard to audio levels that listeners are exposed to in certain listening environments.

The hypothesis of this study was that recordings with a wide dynamic range might be listened to at damaging levels to compensate for the low end of the dynamic spectrum being masked by environmental noise. For example, when listening to music inside a moving passenger car. Experiments were performed to measure the level preferences of drivers in a passenger vehicle to ascertain whether music with a wider dynamic range is listened to at higher levels, compensating for the masked effect at the lower end of the dynamic spectrum. If individuals are listening to dynamic music at a higher average level than compressed music, they are potentially at risk of hearing damage at the high end of the dynamic spectrum.

The results reflect that listeners do not listen to more dynamic music at higher levels than compressed music and it was concluded that the new broadcast loudness standards can also be implemented on material intended for playback in less than optimum listening environments.

Opsomming

’n Ondersoek na Motorbestuurders se Voorkeure in Klankvlakke met betrekking tot Dinamiese en Saamgepersde Musiekopnames

M. Stobbart

Departement Musiek,
Universiteit van Stellenbosch,
Privaatsak X1, Matieland 7602, Suid Afrika.

Tesis: MPhil (Music Technology)

November 2016

Nuwe internasionale wetgewing in televisie- en radio-uitsending en die implementering daarvan deur aanlyn platforms soos YouTube en aanlyn musiekhandelaars soos iTunes is besig om ’n einde te bring aan die dinamies-saamgepersde musiek wat die norm geword het oor die laaste dekade. Alhoewel hierdie nuwe luidheidstandaarde opname-kwaliteit bevoordeel deur ’n groter dinamiese reik toe te laat, mag dit ook onvoorsiene implikasies hê ten opsigte van die klankdrukvlakke waaraan luisteraars blootgestel word in sekere luisteromgewings.

Die hipotese van hierdie studie was dat opnames met ’n wye dinamiese spektrum geluister mag word teen ’n hoë volume om te vergoed vir die onderste deel van die dinamiese spektrum wat gemasker word deur omgewingsgeraas soos as daar na musiek geluister word in ’n bewegende voertuig. Eksperimente is gedoen om die klankvlakvoorkeure van luisteraars in passasiersmotors te meet om te bepaal of musiek met ’n groter dinamiese reik teen hoër klankvlakke beluister word. As daar na meer dinamiese musiek geluister word teen hoër vlakke as saamgepersde musiek bestaan die gevaar dat gehoorskade veroorsaak kan word deur die boonste deel van die dinamiese spektrum.

Daar is bevind dat luisteraars verkies om nie na meer dinamiese musiek teen hoër klankvlakke luister nie. Die gevolgtrekking is dat die nuwe uitsendingluidheidstandaarde ook aangewend kan word in die produksie van materiaal wat gemik is op nie-optimale terugluisteromgewings.

Acknowledgements

I would like to acknowledge the following people as a token of my appreciation:

- My supervisor, Gerhard Roux for his supervision, guidance and enthusiasm regarding my topic.
- My parents, Kent and Carlin Stobbart for their continual love and support, as well as thoroughly proof reading my work.
- My brother, Rowan Stobbart for his active support.
- Andreas van der Merwe, for his creativity in designing the cover art for the printed thesis.
- Samantha Warmerdam, for her constant emotional support and radiant positivity.
- James Trussler, for his guidance in helping me learn LaTeX.
- Will Trim and Jared Prinsloo, for their encouragement and motivation.
- My friends and family for their support and assistance throughout the year.

Contents

List of Figures

List of Tables

Nomenclature

Acronyms and Abbreviations

ATSC	Advanced Television Systems Committee
ARIB	Association of Radio Industries and Businesses
BLW	Between Listener Variability
BS.1770	Broadcast Service 1770
CoG	Centre of Gravity
dB	Decibel
dBA	Decibel Sound Pressure Level, A-weighted
dBTP	Decibel True Peak
dBSPL	Decibel Sound Pressure Level
dBFS	Decibel Full Scale
DR	Dynamic Range
DRT	Dynamic Range Tolerance
EBU	European Broadcasting Union
IEC	International Electrotechnical Commission
ITU	International Telecommunications Union
IRI	International Roughness Index
LFE	Low Frequency Effects

LKFS	Loudness K-weighting Relative to Full Scale
LUFS	Loudness Units Relative to Full Scale
LRA	Loudness Range
LU	Loudness Units
PF	Pink Floyd
PPM	Peak Programme Meter
QPPM	Quasi-Peak Programme Meter
RMS	Root Mean Square
SMPTE	Society of Motion Picture and Television Engineers
SPL	Sound Pressure Level
SRG3	Special Rapporteur Group 3
SS	Saint-Saens
TPL	True Peak Level
VACI	Vehicle Acoustic Comfort Index
VU	Volume Unit
WLV	Within Listener Variability

Chapter 1

Introduction

1.1 Background

To counteract the so-called loudness wars, where recordings and broadcasts have gradually become more dynamically compressed to sound louder than the competition, various broadcast standards based on the ITU BS.1770 (Fleischhacker, 2014:4,5) have been developed. These include ATSC A/85 (USA), EBU R128 (Europe), OP-59 (Australia) and the TR-B32 (Japan) (Fleischhacker, 2014:4). These standards take psychoacoustic principles (ITU-R, 2011:7,8; ITU-R, 2012:10) into account as opposed to the standards of the past, which relied on the measurement of electrical signal levels without corresponding to how audio levels are perceived by humans (Spikofski & Klar, 2004:6).

Whilst the fidelity suffers in over-compressed recordings, due to the distortion being introduced by excessive limiting (Orban & Foti, 2001:2), a reduced dynamic range might be beneficial for listeners in less than ideal listening environments. In this study, a passenger vehicle constitutes a less than ideal listening environment as the lower portion of the dynamic spectrum might be masked by environmental noise.

The hypothesis is that whilst exposed to elevated levels of background noise, subjects will increase the loudness level of each audio track to a level high enough to be enjoyed over the background noise. These damaging sound pressure levels are particularly harmful in dynamic audio and may not however have been the result in compressed recordings, as there is little discrepancy between the high and low levels of the dynamic spectrum.

To test this, an experiment was designed to investigate the preferred comfort level of passenger car drivers, when listening to music over the vehicle's radio system. Participants were asked to listen to two musical tracks, one dynamic and one compressed and asked to adjust the loudness playback level to a level of their enjoyment.

It was found that the mean loudness levels of both the dynamic and compressed musical tracks were less than 80 dBA. Therefore not at high enough loudness levels to potentially induce hearing damage.

1.2 Aim of Research

The aim of this thesis is to investigate the preferred loudness playback level experienced by subjects when listening to different musical tracks over the radio in a passenger car. The preferred loudness levels can be used to determine the desired loudness range of radio listeners, which may aid in the monitoring of loudness normalisation within radio transmission.

The radio environment is simulated through applying audio processing to the wave files of each musical track. The processed wave files are then to be played through the car radio by means of an auxiliary input. The motor vehicle environment is simulated through the use of external speakers producing pink noise at a pre-set level. This gives the simulation of the vehicle being surrounded by city-centre traffic.

In order to help resolve the issue of loudness discomfort, the preferred level of loudness chosen by the subjects within a passenger car will be explored. Reducing the loudness discomfort from compressed audio tracks allows the listener to enjoy the musical transmission over the radio, free from the irritating constant volume adjustments. The preferred loudness values to be selected by the participants will be used to generate decibel values for comparison, measuring the difference in comfort level for each musical track.

1.3 Relevance of Research

This research aims to provide a better understanding of the consequence of inappropriate loudness levels within broadcasted music.

Firstly, as the dynamic range of an audio track undergoes compression, there is a loss of listening pleasure through limiting the dynamic range. In popular music this effect is less noticeable. Conversely, in dynamic dependent music such as classical genres, diminishing the dynamic range results in the loss of delicate notes and the composer's intended dynamic character. The constant production of over-compressed audio across the radio broadcasting industry could appear to the listeners as a wall of sound, with no variation across the musical tracks.

Secondly, the variation in loudness of the broadcasted audio presents a distraction to the driver. With each consecutive track, the driver is prompted to continually adjust the radio's dashboard volume control. This creates an irritating process for the driver as well as a possible hazardous situation.

Thirdly, across the music industry, applications such as iTunes and broadcast websites such as YouTube, are employing the normalisation of LUFS creating a musical shift toward more dynamic music. As the music production makes the shift toward a more dynamic outcome, it is important to understand the use thereof.

Lastly, high levels of loudness exposure remain a long term risk for hearing damage. In order to deal with the constant fluctuations in playback loudness, the volume control may be kept loud to avoid the constant adjustment. With prolonged exposure to higher levels of loudness caused by the lack of normalisation, the listener may be subjected to some hearing damage.

1.4 Structure

Chapter 1: Provides background to the topic, introduces the aims of the study and a brief discussion of the relevance of this research.

Chapter 2: Displays the Literature review across all relevant aspects.

Chapter 3: Describes the Research Methodology as well as the Ethical Considerations and Budget for the study.

Chapter 4: Presents the Results and Analysis for both the Questionnaire and Musical Track Analyses.

Chapter 5: Presents the Discussion relative to similar studies in the field.

Chapter 6: Draws conclusion to the study.

Chapter 7: Presents Recommendations for Further Research and discusses limitations of the present research study.

Chapter 2

Literature Review

The topics affecting the listener's perception include the auditory system; loudness perception; dynamic range and compression; vehicle noise as well as loudness algorithms and descriptors to aid loudness normalisation.

The impact of sustained loud sounds on the human auditory system as well as the possible resulting hearing damage, provide insight and motivation as to why the loudness levels need to be addressed. In order to produce an accurate representation of how each individual is affected by loud sounds, the perception of loudness as well as the dynamic range and compression from a musical stand point needs to be brought to light.

The environment in question is the passenger car where each participant will be seated in the driver's seat. A thorough understanding of where vehicle noise originates from, as well as the impact of traffic noise, may provide insight into the loudness levels at which drivers enjoy music in this environment.

2.1 Auditory System

The sensitivity of the ear and hearing damage are both significant as the outer and middle ear anatomical features impact on an individual's loudness perception. The after-effects of loud audio transmission over radio without the implementation of loudness normalisation may result in hearing damage.

2.1.1 Ear Sensitivity

With reference to the human ear, the outer ears and ear canals are individually different in shape, size and structure, whilst the auditory system functions similarly in all individuals. The subtle individual structural differences can significantly influence the perception of loudness. Nocross & Thibault (2011:3) highlight that loudness perception is influenced by the frequencies and intensity of the incoming sound, as well as the listener's unique auditory structure. The human ear is able to perceive the faintest sounds as well as the most intense, including sound pressure levels up to 120 dB, without sustaining permanent hearing damage.

The human ear is most sensitive to the middle frequency range sounds between 1 kHz - 5 kHz (Nocross & Thibault, 2011:3; Plack, 2004:5), specifically around 3kHz as this is closer to the resonant frequency of the auditory canal (Nygren, 2009:4). The higher and lower frequencies outside of the sensitive range need a higher intensity output in order to be perceived by the ear at an equal loudness level to the middle-frequencies. This equal loudness curve constitutes the loudness contours of each individual frequency in relation to each other and is called the Fletcher Munson Curve (fig. 2.1).

The data presented in Figure 2.1 originates from an amalgamation of research by Steinberg and Fletcher. During the years 1921 - 1924, their research comprised of measuring loudness by presenting stimuli which exceeded some threshold by a certain number of decibels, coupled with a formula to calculate the loudness of any complex sound. Their research was reviewed by Bell Telephone Laboratories resulting in the 1933 paper researching experimental methods for calculating loudness of complex sounds (Fletcher & Munson, 1933:82,83).

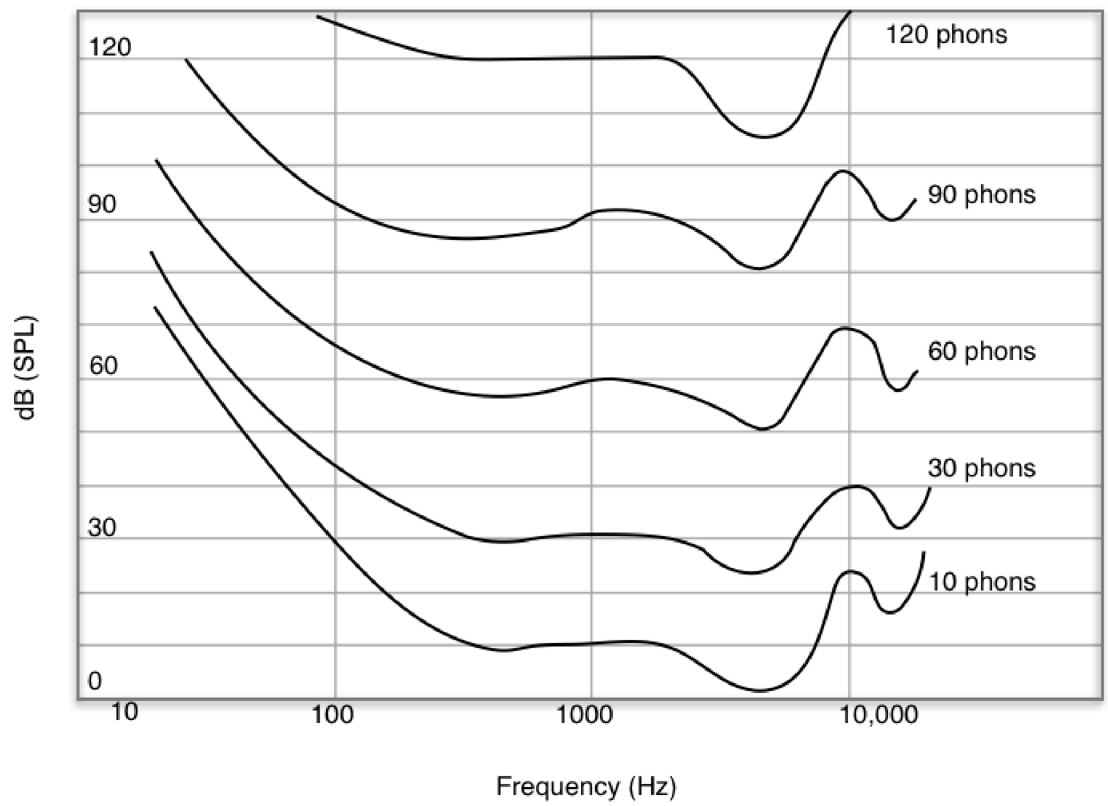


Figure 2.1: Loudness Contours (Fletcher & Munson, 1933:91)

The loudness contours show the amplitude level at which a sound must be produced at in order to be perceived equally as loud as a 1 kHz reference tone. Due to the low resonance of the bass frequencies, the dB SPL has to be much higher to be perceived at equal loudness. Conversely, the higher frequencies can be produced at a lower dB SPL. In order for 10 kHz measured at 30 phons to be perceived equally as loud as the 1 kHz reference, it has to be produced at 40 dB SPL.

The amplitude rating on the loudness contours relates to how an individual's ears interpret and transduce incoming sounds. The graph accurately portrays individualised loudness perception as well as the complexity associated with developing a singular method for loudness normalisation. As each individual's head and ear shapes are different, the loudness contours will differ along the curves. Therefore, developing one algorithm that would comfortably fit the demographic for all the listeners would prove to be exceedingly complicated.

To portray how detrimental high intensity incoming sounds can be to an individual's auditory sensitivity, the hearing and pain thresholds have been illustrated using sound pressure levels (SPL). This means that at 0 dB SPL, a incoming sound is a barely audible, whilst 130 dB SPL is the threshold whereby most individuals will experience pain from the incoming sound (Howard & Angus, 2009:92; Davis & Brown, 2013:97).

According to Fleischer (2008:112), the threshold of pain is described as when

> "loud sound is tearing at nerve endings that signal the impression of pain. Such nerve-endings are said to be in the tympanic membrane, as well as in the joints and ligaments of the middle ear".

Therefore, the individual's experience of the pain threshold results in mechanical damage of the ossicular chain located within the middle ear.

When exposed to high intensity sounds, in excess of 80 - 100 dB SPL, the threshold of pain reduces. Therefore, when the individual is exposed again to heightened noise levels, pain will be experienced at a lower dB SPL (fig. 2.2). It is important to be aware of the pain thresholds because the music volume over radio broadcast currently transmits compressed audio with a large volume variation, which, may affect the hearing of listeners.

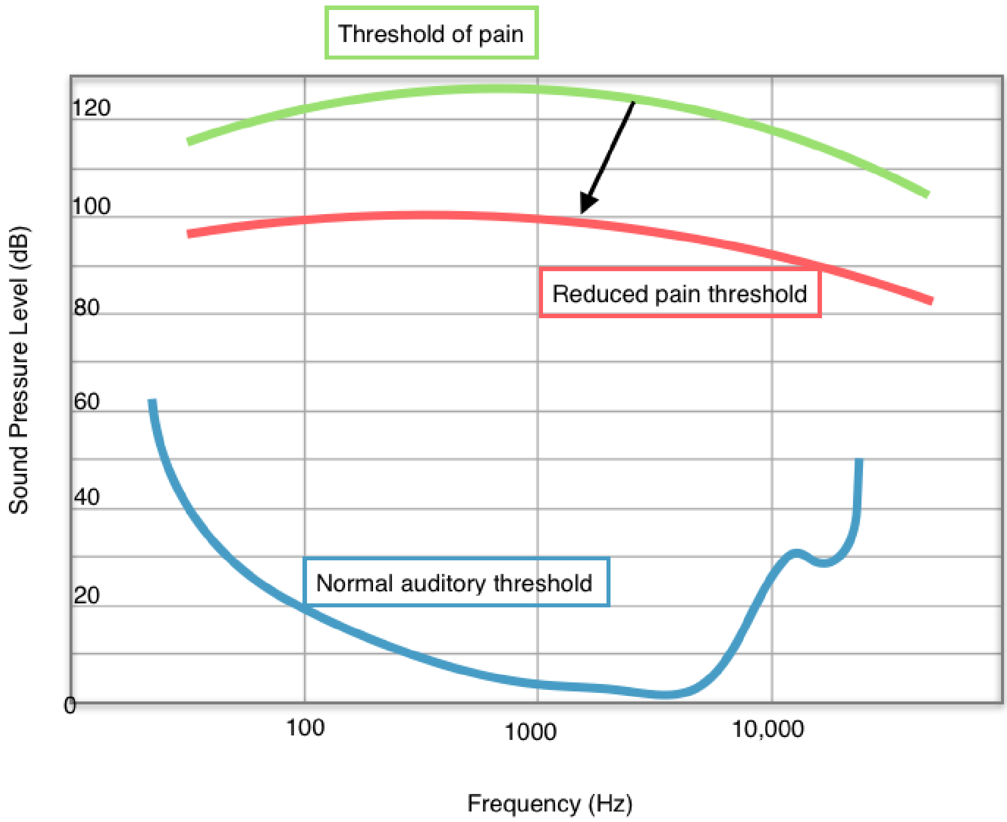


Figure 2.2: Lowering Pain Threshold (Fleischer, 2008:113)

The acoustic environment within which the individual perceives incoming sounds can greatly influence their perception of the pain threshold. The passenger car environment does not always provide the most ideal or acoustically treated environment for musical enjoyment. Therefore, fluctuation in background noise means that listeners are subjected to ever louder sounds from both inside and outside the vehicle. The interior noise levels may be attributed to speech and radio loudness, whereas Ouis (2001:105) highlights that tires, engine and exhaust, coupled with the air turbulence contribute to the heightened exterior noise levels.

The International Telecommunications Union (ITU) and the European Broadcasting Union (EBU) aim to normalise the loudness variation problem. Currently intruding traffic reports or adverts, coupled with the surrounding background noise may still peak high enough to cause pain to the listener.

With the change in environment, the surrounding background noise varies as shown in Table 2.1. According to Davis & Brown (2013:97), the typical dBA reading for light traffic measured at 30 m is 50 dBA, whereas a sports car travelling at 90 km/h will have a background noise level of 80 dBA. Environmental SPL values provide assistance with the choice in loudness level for the background noise simulation used in this experiment.

Table 2.1: Environmental Sound Press Levels (Howard & Angus, 2009:92)

Environment Summary	dB(SPL)	Explanation
Close up gunshot	140	
Threshold of pain	130	Painfully loud!
Jet take-off	120	
Night Club	110	
Aggressive Shouting	100	Very noisy
Large truck	90	
Heavy traffic	80	
Passenger car interior	70	Noisy
Regular conversation	60	
Office environment	50	
Living room	40	Soft
Bedroom at night	30	
Vacant concert hall	20	
Calm breeze	10	Just audible
Threshold of hearing	0	

In order to determine the best interior musical broadcast level within a passenger car, prior knowledge displaying that heavy traffic noise levels may reach 80 dB SPL, will ultimately help provide guidance as to the ideal background noise loudness level for the experiment. The background noise level needs to be high enough to compensate for the absorption and reflection caused by the vehicle. This allows for the residual noise that bleeds into the vehicle to be measured.

2.1.2 Hearing Damage

Even though the broadcasting industry is attempting to overcome the problem of audio loudness compression, many drivers are at risk of hearing damage, or at least, exposed to auditory discomfort during musical playlists, news interjects and adverts.

As a result of musical loudness, drivers themselves are subjected to a temporary threshold shift affecting their hearing. According to Skovenborg & Nielsen (2004:3), this temporary shift, which can last several hours, occurs after exposure to loud sounds causing a reduction in hearing sensitivity. The effect of a temporary threshold shift may still present a hazardous situation particularly within a vehicle. A loss in hearing sensitivity results in a reduction of the nerve impulse's efficiency to transmit incoming sounds. Moreover, a loss in acuity means an individual's mechanism of positive feedback on the enhancement of standing waves within the cochlea becomes damaged, hindering the ability to accurately distinguish between incoming sounds.

Should the loudness normalisation be left unaddressed, more people may be at risk for developing tinnitus, a condition whereby the cochlea spontaneously produces noise in the form of tonal or random noises (Howard & Angus, 2009:102).

In an attempt to contain the playback of very loud music, the European Legislation reduced the level twice - first to 85 dBA and then to 80 dBA. The aim of this reduction was to regulate loudness levels within a noisy work environment. If workers are subjected to loudness levels higher than the first step of 80 dBA, the employers are required to provide hearing protection for the employees (Howard & Angus, 2009:104).

It therefore seems surprising that if employers are required by law to provide hearing protection to employees working within an noisy environment in excess of 80 dBA, then individuals simply enjoying music are subjected to uncontrolled loud interjects which may peak above 80 dBA. According to MusicLoudnessAlliance (2012:3), several European countries have attempted to address the issue of developing hearing damage by limiting peak level loudness output. The results however, have revealed difficulties in listening and appreciating classical as well as other genres at reasonable loudness output, without exceeding legislative guidelines.

Instead of resolving the loudness issue, the legislation placed more emphasis on the mastering engineers to reduce the dynamic properties of said genres to be enjoyed within noisier environments (MusicLoudnessAlliance, 2012:3). Therefore, the audio quality broadcasted to the audience is vastly reduced in order to allow all genres to be transmitted into any environment the audience may be listening in.

2.2 Loudness Perception and Subjectivity

The variables relating to the perception of loudness and subjectivity experienced by individual listeners tuning into radio broadcast are detailed below.

2.2.1 Loudness Wars

Loudness refers to the individual perception of audio intensity through a playback medium such as radio broadcasting. The use of dynamic compression has caused what is commonly known as the loudness wars. According to Apple (2012:4), this divides the music industry into the artists and producers who feel ever increasing loudness is better and the audiophiles who argue increasing loudness diminishes dynamics and headroom. This divide impacts significantly on the entire music industry with implications for the quality of sound listeners are exposed to. The present investigation will focus on radio broadcasting and the impact on loudness preferences of listeners in a vehicle environment.

Radio stations process each musical track to increase the output loudness. This generates a so-called loudness war between broadcasters. The so-called loudness war dates back to the beginning of recorded music where it became well known due to Phil Spector's role in mixing. Southall (2006:1) highlights that Phil Spector's production style, mixing and mastering aimed to cram as much sound as possible into a small space. The term *Wall of Sound* was used to describe his method of generating vinyl track loudness. According to Vickers (2010:2,3) Phil Spector's use of the *Wall of Sound* used an echo-chamber to stimulate the peak amplitude resulting in a higher RMS power. This was achieved through natural reverberation and large ensembles creating increasingly loud music. Through the integration of the *Wall of Sound* production style, the vinyl track loudness levels increased creating competitiveness amongst record sales.

The competitiveness between record sales has led to an ever growing competition between producers and broadcasters. The so-called loudness wars developed over radio broadcast, where each station attempted to be louder than their competitor. Orban & Foti (2001:1,2) state that in 1998, compact discs (CD)s used in broadcasting had pre-distortion processing and intentional clipping in attempt to increase their overall loudness. Through the use of phase rotations by radio processors, the overall on-air clipping of the audio track would not increase. This affected the quality of sound experienced by the listener by keeping it at a more consistent loudness level. Unfortunately, as broadcasting stations compete for the loudest on-air sound, the listeners become prone to perceiving the changes in tracks and the crossing between stations, as too loud. One listener may find the popular rock music on the radio to be at a reasonable level, the same listener may find the classical rock music on the opposing radio too loud. This means that the listener is required to adjust his/her volume control with each switch between stations.

2.2.2 Loudness Perception and Preferences

The perception of loudness is highly subjective, which according to (Wolters & Riedmiller, 2010:4) involves physiological and psychoacoustic factors unique to each individual listener. Due to the individuality of the pinnae, coupled with the subjectivity of perception, the creation of a single measurement that works universally, is therefore complicated.

Fletcher & Munson (1933:82) portray loudness as a psychological term describing the auditory sensation of magnitude. The use of *pp*, *p*, *mf*, and *ff* to describe whether a sound is perceived to be soft or loud give a limited description as the understanding depends on the experience of the listener. Should the listener have a musical background, their understanding of how loud a sound is perceived from pianissimo to fortissimo will be more accurate than a regular untrained individual's response. In order to determine loudness, it is important to observe and define the sound's intensity, physical composition as well as the physiological and psychological conditions of the listener. The psychological conditions as pointed out by Fletcher & Munson (1933:82) include the emotional state, alertness, fatigue and attention span which also affect the response and perception. It is therefore important to be aware of the psychological factors as it could impact the loudness perception result as each listener may be in a different emotional state.

In support of Fletcher & Munson (1933:82), Lund (2006:59) explains that the subjectivity of loudness is perceived differently through the SPL, frequency contents and the duration of the sound. This makes it impossible for each listener to perceive the same sound or musical track in the same manner. Furthermore, it is important to define the loudness measurement in terms of the listener demographic. The loudness definition will concern Between Listener Variability (BLV), which relates to the differences in perceptions of a group of people as well as their culture, age and sexual orientation. Another loudness differentiation used by Lund (2006:59) is Within Listeners Variability (WLV) which refers to the change of timing, mood and attention of a singular participant. For this thesis the focus is more on the BLV and therefore the WLV will not be explored any further.

The subjectivity of loudness justifies a zone whereby listeners feel that the volume is at a reasonable playback level. This, according to Riedmiller & Robinson (2003:3) implies a loudness range satisfying the listener's volume preference for musical playback. This is referred to as the comfort zone (fig. 2.3).

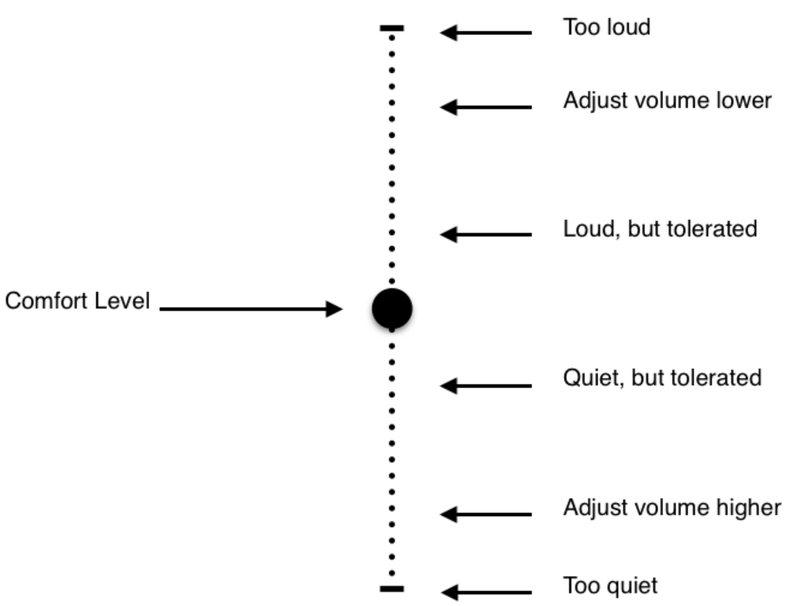


Figure 2.3: Listener's Comfort Zone (Riedmiller & Robinson, 2003:5)

The notion of creating the perfect zone that satisfies all listeners is virtually impossible, as every individual listener has a different opinion on what is considered to be too soft, too loud or just right. In addition to creating the ideal loudness level for listeners, the environment in which the perception takes place greatly impacts the listener's comfort zone. For example, music broadcasted in both shopping malls and cars have to compete with the elevation of background noise in order to be perceived clearly. For some listeners the background noise could cause a distraction and thus in order for the music to be properly enjoyed, the volume of the music would need to be elevated much higher than the surrounding noise.

The genre of the broadcasted music plays a role in determining what is considered to be the most comfortable level for auditory playback. Riedmiller & Robinson (2003:5) state that a "rock concert [...] would seem silly if they were not louder that a current affairs discussion" . This means that equal loudness across musical genres is not desirable.

Skovenborg & Nielsen (2004:1) highlights that loudness perception not only depends on the volume, but also on the format in which the track is played. Radio broadcasting as well as music on CDs, undergo spectral processing in order to make the music more aesthetic (Skovenborg & Nielsen, 2004:1). As a result, the listener may experience jumps between audio tracks or between radio stations. In order to eliminate the variability of the audio format, this thesis will ensure that all audio files played utilise the same audio format. Using the highest quality audio source alongside exporting all the tracks to -23 LUFS, the chance of a harmful audio spikes between tracks will be eliminated.

2.3 Audio Processing

This section will explain how dynamic range, hypercompression as well as distortion and masking, affect the loudness output of musical tracks.

2.3.1 Dynamic Range

It is understood that radio stations attract listeners through sound appearance and musical preferences. Maempel & Gawlik (2009:1,2) highlight that the goal of radio stations is to impress the audience with a unique sound through positive attributes. This is done through a number of relevant perceptual criteria, including: the music's aesthetic impression; track recognition; listening convenience; intelligibility as well as brand value communication. Each musical radio station will exhibit any number of these qualities in order to ensure the long-term interest of the listening audience.

It is understood that music produced for transmission has reached a point of constant peaking levels across the audio waveforms. This means there is no longer a distinct difference between the high and low amplitude sections of musical tracks because the production has resorted to compressing the dynamic range. With the reduction in dynamic range through audio processing, it is commonly felt that "much of the music we listen to today is nothing more than distortion with a beat"

(Speer, 2001). Therefore, to combat the poor musical quality, the broadcasting industry is attempting to monitor these processes to deliver a better quality of audio for future transmission.

Wolters & Riedmiller (2010:1) state that the film industry was first to deal with the complication of varied mixing and loudness output through the integration of a collection of worldwide recommendations. These recommendations are developed and monitored by the Society of motion Picture and Television Engineers (SMPTE) (SMPTE, 2016). This allows for loudness control to be regulated across all theatres. The recommendations that govern the loudness control within motion picture theatres would be ideal for radio broadcast, specifically between tracks, adverts, interjects and the imminent crossing between radio stations. The importance of loudness normalisation across radio transmissions will ensure an overall increase in quality and audibility of sound appearance, which, according to Maempel & Gawlik (2009:1) is the main objective that each radio station strives to provide for their listeners.

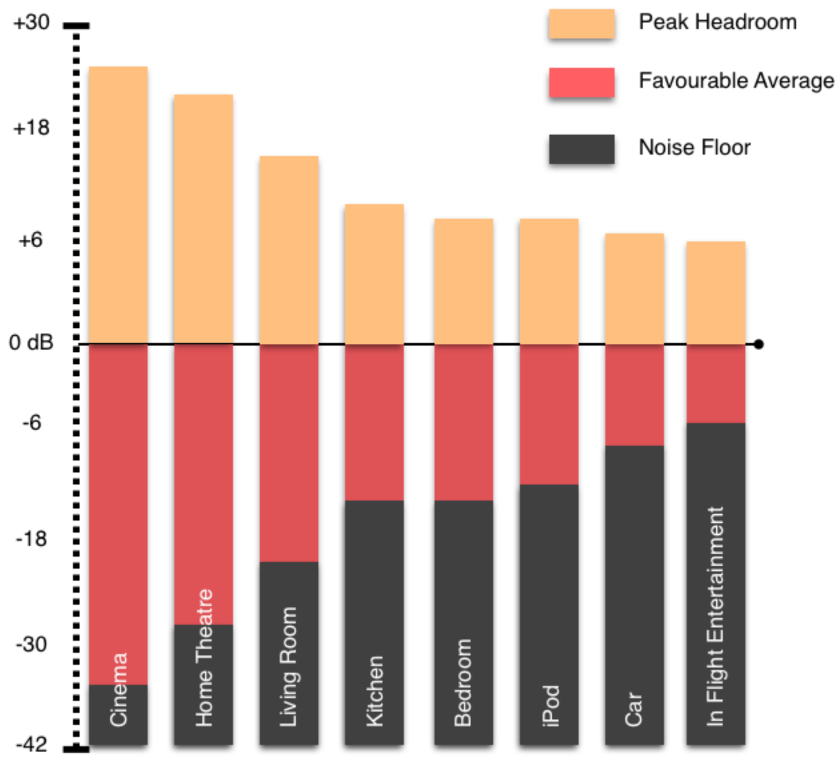


Figure 2.4: Environmental Dynamic Ranges (Hadi, 2010:10; Lund, 2006:57)

The optimum dynamic range varies with the listening environment (fig. 2.4), as more dynamic genres would require a lower noise floor in order to be heard as intended by the composer. This is notable because radio broadcast can be mixed specifically for an environment with a higher noise floor, meaning the listener will raise the volume in order to hear the broadcast. Hadi (2010:10) highlights that having a wider dynamic range for musical broadcast is ideal, however the effectiveness of the dynamic range perception depends on the noise floor of the listening environment.

Lund (2006:57,58) describes a phenomenon called Dynamic Range Tolerance (DRT) which refers to the favourable average window plus the peak level headroom for each musical track. The DRT, does however depend on the listening environment. Within the vehicle environment (fig. 2.4), there is a higher noise floor and lower headroom which means the DRT is smaller. In contrast, music played within a living room has a lower noise floor and more headroom meaning the DRT value will be greater. To ensure a decent signal to noise ratio and DRT, music played within a living room can be set at 45 dBA SPL, whereas the music within a vehicle can be at 65 dBA SPL (Lund, 2006:58).

A similar description of the DRT is given by Skovenborg & Lund (2009:8) in the form of an individual's listening tolerance. Here the description of the DRT is stated as, "the typical distance between RMS level and peak level that a consumer would tolerate inside a programme or musical track" (Skovenborg & Lund, 2009:8). The DRT works similarly to the Comfort Zone (Riedmiller & Robinson, 2003:3) phenomenon whereby the loudness levels reached by the audio content outside of the listeners' comfort zone would create irritation, annoyance and an urge to turn the audio content down. The problem with the DRT and Comfort Zones is that both are asymmetrical and entirely subjective to the circumstances of the listener.

In addition, Speer (2001) states that music used for radio broadcast undergoes further processing to make the track *radio ready*. This is a term coined by marketing professionals who use music with the intention to sell a product or service. This means that the soft and more dynamic tracks are raised to a level that forces them to compete with naturally loud tracks, resulting in a vastly reduced dynamic range amongst the softer tracks.

Prior to radio transmission, every musical track goes through audio processing in addition to the mixing and mastering. Orban & Foti (2001:1) have indicated that audio processing functions, such as a series of limiters, are used to control the peak modulation and ensure that the track meets legal requirements. The limiters reduce the peak-to-average ratio significantly, allowing the radio station to give the illusion of being louder within the allowed peak modulation limits.

From a dynamic range compression perspective, Nielsen & Lund (2003:5,6) discuss a system of identification, utilising a short 0 to 4 rating of *hotness* to allow individuals to assess whether CD albums retain a good dynamic range. For example: remastered Oye Como Va by Santana 1970 - 1999 has a rating of 1, which means the track is well balanced using the full range of the CD. Conversely, Smooth by Santana, released in 1999 has a rating of 4. This means the track contains too much dynamic processing and distortion, giving it a hot rating. Having a rating system for the focus of high fidelity audio is valuable, especially to radio broadcasters.

Lund (2006:59) states that classical and talk radio stations currently utilise high fidelity audio. The ideal listening environment for different musical genres affects the appropriate dynamic range requirements for the music to be enjoyed. For example, the best environment for classical music would be with low background noise. This will allow the listener to enjoy the full expression of the classical recording.

Since classical music recordings have a greater dynamic range and therefore a lower *hotness* rating, the vehicle environment with heightened background noise is not ideal for optimal listening of classical genres. In order for popular and rock music to be suited for a vehicle environment, radio stations utilise a hotter rating for each track to combat the loud background noise. The audio processes responsible for raising the hotness rating for music tracks are hypercompression, distortion and masking.

During this investigation, the hotness scale was used to ensure each musical track had a low hotness rating, thus allowing for the best unprocessed tracks for the experiment. The process of hotness measurement is discussed in the Loudness Integration section.

2.3.2 Hypercompression

Listeners described in this document will be exposed to musical tracks and therefore it is relevant to consider the impact of hypercompression, especially on popular music. When listening to musical tracks on the radio, tracks exhibit degrees of audio processing which include hypercompression, distortion and masking which, if used in excess, can be detrimental to audio quality.

Within the music industry, an advocacy group, Music Loudness Alliance[1], highlights that sound quality reduction across musical production is caused by the peak normalisation of audio tracks. In addition to the peak normalisation affecting the dynamic range, hypercompression and music clutter also cause audio quality damage through the reduction of musical emotion, punch and clarity (Speer, 2001; Vickers, 2011:346). The problem persists when each musical track is compressed further than the last, subjecting listeners to increasingly louder sounds.

Hypercompression is the result of audio processes used by producers, when they attempt to add more loudness and density to their musical tracks. Apple (2012:4)'s documentation states that within the music industry, artists and producers disagree how compression and mastering should be implemented, as

> "some feel that overly loud mastering ruins music by not giving it room to breathe, others feel that the aesthetic of loudness can be an appropriate artistic choice for particular songs".

[1] According to a white paper released by the group, Music Loudness Alliance consists of leading technical and production members. These include professionals Eelco Grimm, Kevin Gross, Bob Katz, Bob Ludwig and Thomas Lund, led by Florian Camerer (MusicLoudnessAlliance, 2012:1)

It is understood that the audiophiles prefer music with the original dynamic range, unprocessed and as the artist intended, whereas broadcasters and producers make use of compressed audio for a competitive advantage over the other radio stations within the same genre.

Vickers (2010:1,2) defines hypercompression as the squeezing of more loudness into a recording no matter the consequence to audio quality, which, when coupled with the overuse of mastering for loudness, results in the deterioration of musical quality. Through hypercompression of digital audio, the musical quality suffers at the reduction of the dynamic range, destroying musical emotion (Levine, 2007).

The damaging effect of hypercompression (fig. 2.5), shows the difference between the normal, un-compressed waveform and the hypercompressed waveform with reduced dynamics and thus less musical emotion. The tracks were processed with hypercompression in Logic Pro (fig. 2.5).

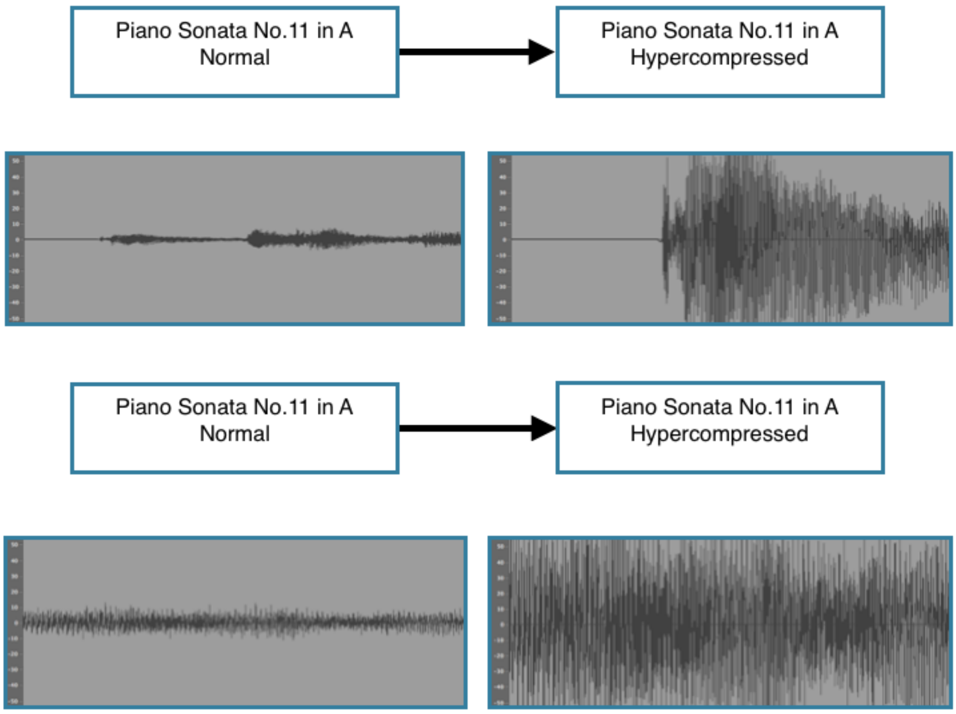


Figure 2.5: Hypercompression on Classical Music

The musical tracks used within this experiment have not been subjected to hypercompression, in order to keep the tracks with a dynamic range as large as possible. Instead, the tracks used in the experiment have been normalised in accordance with new broadcasting standards to preserve dynamic quality. It is important to acknowledge the detriment effect of hypercompression as current musical tracks used in radio transmission are heavily compromised by it.

Orban & Foti (2001:3) point out that once the music is subjected to hypercompression as well as the required marketing amplitude levels, the resulting transmission portrays lifelessness and an overall lack of drama. Similarly, Southall (2006:1) highlights that music should be perceived without hypercompression as,

> "music isn't meant to be at a constant volume and flat frequency; it's meant to be dynamic, to move to fall and rise and to take you with it, physically and emotionally".

Due to hypercompression, the notion of whether louder is better becomes contradictory as Orban & Foti (2001:4) allude to the fact that radio broadcast directors aim to have their music loud constantly. The persistent loudness reduces the risk of listeners skipping over the station whilst tuning the radio, or assuming that the receiving signal is weak and thus unsatisfactory. It is with this mindset that hypercompression needs to be re-evaluated for the production and radio broadcast of music. In this investigation, both popular and classical musical tracks are used to ensure diverse musical stimuli. The popular track is especially important as it embraces a moderate dynamic range free from added hypercompression, therefore resembling how it should be broadcasted over radio. The classical track embraces a large dynamic range also free from detrimental audio processing providing a glimpse of how classical tracks are broadcasted over radio.

The environment and attentive state in which the listener perceives music also has an impact on whether the music needs to be dynamically processed to be enjoyed. Rogers (2011:10) states that music is generally enjoyed over three separate locations. These include: a single room, in a vehicle or house, as well as portable music players used within an ever changing environment. This is important to note, as the background noise in each location is vastly different and the amount of audio processing required should therefore support the environment in which the music is played.

For a live performance in a single room with minimal background noise, Rogers (2011:10) states that the attention of the listener is focussed solely on the performance and thus the best dynamic version of the music should be heard to give the best experience. In the case of a vehicle, where background noise is constant, the attention of the driver is focussed on the road and surroundings rather than directly on the music, meaning that the addition of some compression via the radio processors is ideal to raise the signal to noise ratio.

2.3.3 Distortion and Masking

In addition to hypercompression, distortion and masking are also detrimental to broadcasted audio. Rogers (2011:4) highlights that harmonic distortion has both positive and negative effects. The positive effects, when used in moderation, can alter the *timbre* of the music for a desired effect, however when coupled with hypercompression, the listeners are quickly subjected to listening fatigue.

Within a vehicle, where a listener may be exposed to radio interjects coupled with loud music, warning sounds along the road can be masked. This could result in a driver becoming unaware of possible dangers. Fleischer (2008:91) points out that low frequencies can suppress the perception of high frequencies. However, this suppression happens over the critical bands rather than over the full frequency spectrum. The suppression of even part of the upper frequency spectrum may hinder the listener from receiving critical information about the location of possible dangers. For the environment of the driver, not hearing the high frequency sounds for example cars hooting, may have dangerous consequences. The suppression of high frequency perception depends heavily on the masking tone's amplitude, which, according to Fleischer (2008:91) will reduce the masking effect or even eliminate it completely. With a vehicle in motion, the low frequency rumble from the engine as well as the constant traffic noise create enough of a masking effect prior to the inclusion of hypercompressed audio.

2.4 Vehicle and Background Noise

This section will present an overview of the noise levels experienced by drivers. Awareness of noise factors are important as they can significantly affect the driver's comfort when listening to music in motion.

2.4.1 Traffic Noise

According to Ouis (2001:106), the noise surrounding the vehicle can be broken down into the individual vehicle's sound emittance and the collective noise from surrounding road traffic.

Each individual car emitting sound, acts differently when compared to the collection of travelling vehicles. Ouis (2001:106) highlights that sounds created by individual cars diminish by 6 dB as stated by the inverse square law, where the listener doubles the distance between them and source of noise. Conversely, a collective group of travelling cars all producing noise, creates a consistent soundscape, which can be described as background noise.

In a study conducted by Lewis (1973:193), a variety of vehicles including cars, vans, trucks and buses were tested for their output noise. The study revealed that the majority of vehicle background noise is perceived at approximately 80 dBA. The heavier vehicles generate a louder sound level (80 - 85 dBA), whilst the lighter vehicles and those on the opposite side of the road produced a significantly lower overall sound level (65 - 75 dBA). The vehicle to be used in this investigation is a Renault Clio 2006 model, which, as a lighter vehicle falls below the 80 dBA output level.

For the purposes of this investigation, the noise levels surrounding the vehicle will be in the form of pink noise[2] played at a predetermined level through two active speakers providing a simulation of the traffic noise. The simulation of traffic noise eliminates the independent noise variables present whilst driving a set route at a set speed. The simulated traffic noise ensures that each participant is exposed to the same noise output.

In addition to the noise generated by the traffic, Nor & Ariffin (2008:344) state that vibrational noise also contributes to background noise. Vibration sources include the suspension of the vehicle, the driver's travelling speed and the roughness of the road. The knowledge of the parameters within the International Roughness Index (IRI) with the addition of information from the Vehicle Acoustic Comfort Index (VACI) may help design the ideal comfort environment for drivers to enjoy listening to music while driving (Nor & Ariffin, 2008:344,345).

The focus of this study is on the preferences in loudness pertaining to the musical experience and not on the external and mechanical features of the car. Therefore, the external noise factors have been dealt with through a simulation using pink noise instead of a detailed analysis of each vibrational noise source. The suspension of the car, choice of speed and road roughness do not play a role as the simulation is set with the car at rest, therefore not generating any friction that would add to the background noise.

The interior reference level set for the experiment is at 60 dBA. This value was calculated by driving around Stellenbosch using the Mic-Wi436 in conjunction with DSP Mobile Analyzer software displaying the L_{eq} of the interior noise levels. The car was driven at speeds up to 80 km/h on the highway and at an appropriate level within the city-limit. Furthermore, the noise level reference was compared with similar research to ensure a precise value was selected. The reference level of 60 dBA was decided as it best simulates the bleed of background noise into the vehicle. Therefore, the background noise is sufficiently loud to simulate that of a city-centre driving environment.

To provide an interior noise level of 60 dBA, the exterior noise reading should be set at a higher level to compensate for the reflection and absorption of the car's framework. It can be seen in Table 2.2 that during an experiment carried out by Bjorkman & Rylander (1997:514), the discovered noise output from an array of vehicles reveal that the majority do not exceed 75 dBA.

Table 2.2 below shows the number of vehicles (n) measured per vehicle type, giving a percentage distribution (%) of vehicles above and below 75 dBA. The vehicle types range from a small passenger car through to larger transport trucks.

---

2 Pink Noise is created through an equal energy distribution over each octave (AcousticFields, 2016)

Table 2.2: Vehicle Noise Output (Bjorkman & Rylander, 1997:514)

Type:	Passenger Vehicle		Van		Medium truck		Bus		Large truck	
dB(A)	n	%	n	%	n	%	n	%	n	%
<75	120	94	35	97	27	54	23	88	14	93
>75	7	6	1	3	23	46	3	12	3	7

From Table 2.2, it can be seen that 94% of the passenger vehicles measured in (Bjorkman & Rylander, 1997:514,516)'s experiment do not peak above the maximum value guideline of 75 dBA.

Similarly to the noise levels recorded by Bjorkman & Rylander (1997:514), it can be seen that the difference in noise levels measured on both highways and in city-centres reveals a different result. Ouis (2001:107) states that the city-centre noise is more asymmetrical than highway noise displaying 60 - 80 dBA in variation, whereas the highway noise displays a more sturdy range of 70 - 80 dBA.

The vehicle noise values from Bjorkman & Rylander (1997:514) and the environmental values from Ouis (2001:107) were compared with the personal observations whilst driving both in the city-limit and on the highway. This helped to determine the ideal background noise level for the experiment.

In this investigation, the reference level within the vehicle was attained by playing pink noise at a varied levels until the bleeding noise into interior reached the predetermined level of 60 dBA. Therefore the background noise level would be positioned between 60 - 80 dBA, simulating the surroundings of a city centre.

2.4.2 Airborne and Structural Noise

An understanding of airborne and structural noises experienced by drivers, give a more comprehensive foundation to the noise exposure that would affect drivers' comfort when listening to music over the radio. The values presented in this section provide the researcher with background knowledge on the decibel levels drivers experience whilst driving. This in turn allows the researcher to define a more accurate preset level to broadcast pink noise to the vehicle and driver within the controlled environment.

In order to create the most comfortable within vehicle listening experience for the drivers and passengers, the relationship between noise levels and musical discomfort must be understood. The interior atmosphere that can be experienced by passengers is affected both by the make and brand of the vehicle as some will have more isolation from the outside environment. Ormuz & Muftic (2004:77) describe the distinction between the feelings of comfort and discomfort with reference to the well-being of the individual, where

> "comfort implies a conscious well-being. Discomfort implies a consciousness of unwell-being, corresponding to feelings such as annoyance or irritation".

It is therefore understood that creating the perfect listening environment deemed comfortable by everyone is in fact impossible.

The airborne and structural noises affecting the driver allows for the researcher to design a more precise simulation of the background noise. This is done through analysing the interior musical experience and how it is affected by the comfort zone of each participant. The indoor residual noise level will be used as a reference as it takes into account the insulation of the driver from the background noise. This reference value will be used against the loudness level of the music adjusted by each participant to show how much higher each participant requires their comfort zone. Due to the individuality of each participant, this loudness comparison will provide a wider array of preferred musical listening comfort levels.

More information pertaining to the integration and use of background noise within the controlled environment is presented in Chapter 3. The following section will detail the different loudness algorithms that have been developed to ensure the reduction and normalisation of audio spikes within the broadcasting industry. It is through the implementation of these algorithms that the comfort zone and enjoyment of music by each driver can be improved.

2.5 Loudness Algorithms

It is necessary to highlight the benefits and shortcomings of each loudness normalisation algorithm in order to understand their contribution to the broadcasting industry.

For this investigation, understanding the EBU R128 and it's development is mandatory as it shows the progression into loudness normalisation across the industry. This experimental procedure will utilise a long term loudness output value of -23 LUFS, in accordance with the EBU R128.

2.5.1 ITU-R BS. 1770

As it stands within the radio broadcasting industry, there are channel to channel loudness level discrepancies proving that the present track normalisation methods are severely lacking (Riedmiller & Robinson, 2003:1). Progress is however being made through the standardisation of algorithms to monitor track normalisation and spikes in audio levels. The management of loudness normalisation within radio transmission is especially challenging due to the audio content fluctuations which, according to Soulodre & Lavoie (2005:1) include the constant interchange between music, speech and a combination of sound effects. Therefore, the development of a single standardisation to normalise these audio fluctuations proves to be a complicated procedure.

The Special Rapporteur Group (SRG3) situated within the ITU developed an objective metering system to measure perceived loudness of audio programme materials within the broadcasting sector (Soulodre & Lavoie, 2005:2). This lead to the development of a loudness normalisation algorithm with the intention of reducing volume variations between musical tracks and broadcasting stations. According to Robjohns (2014:114) the initial release of the ITU-R algorithm was in September 2007, titled the ITU-R BS.1770. Since its release, the algorithm has undergone several revisions up to the current revision as of October 2015, the ITU-R BS.1770-4 (ITU-R, 2017:1).

This basic loudness measurement developed by the ITU forms the foundation of most loudness normalisation algorithms. Robjohns (2014:114) highlights that the ITU-R BS.1770-3 includes four distinctive stages to accurately measure subjective loudness. These include: response filtering, average power calculation, channel weighting and summation as shown in Figure 2.6.

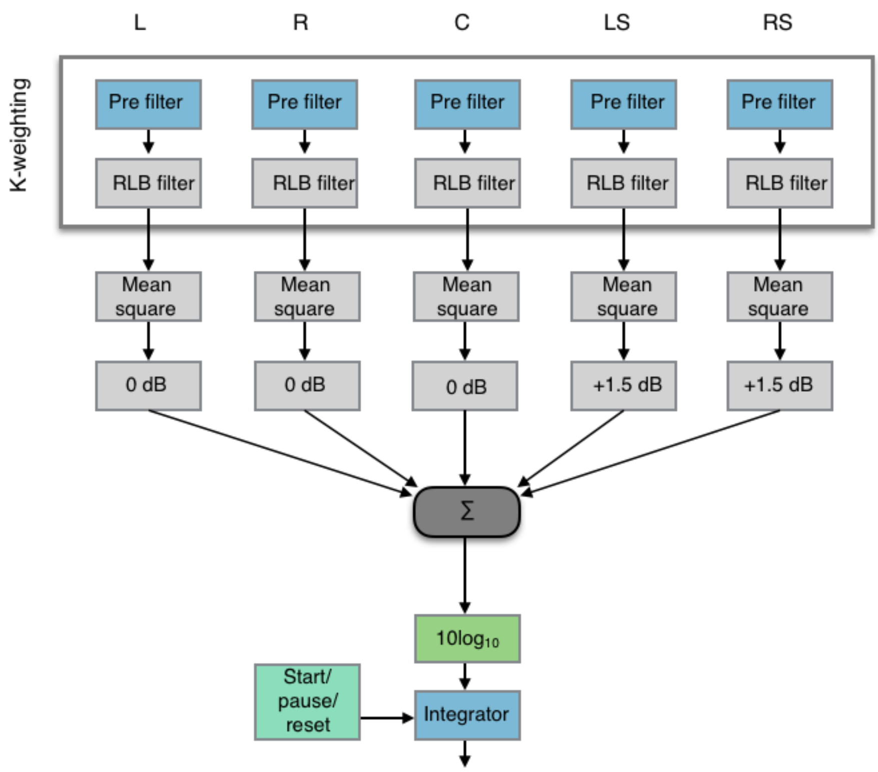


Figure 2.6: Channel Processing of the ITU-R BS.1770 (Camerer, 2010:3)

The ITU-R BS.1770 algorithm (fig. 2.6) operates through the use of pre-detection filters, RMS measurements for each audio channel as well as the summation of the channel powers (Adriaensen, 2011:12). Each of the audio channels undergoes individual filtering with a low frequency roll-off and a high frequency shelf. This filtering effect, according to Cabot & Dennis (2011:2), is titled K-weighting and simulates the sensitivity of the human ear as well as head diffraction effects. Similarly, Fleischhacker (2014:6) states that K-weighting filtering is designed to emulate the acoustic effects of the human head. Camerer (2010:3) points out the importance of the K-weighting curve, as it creates the foundation with which the inherent subjective impression and objective measurements of a given sound can be matched.

Moreover, the K-weighting curve is applied to all the channels with the exception of the Low Frequency Effects (LFE) as it is discarded from the measurement. The signal then undergoes the RMS calculation before the final result being produced in the form of Loudness K-weighting, relative to Full Scale (LKFS)(Camerer, 2010:3). The surround sound channels, as highlighted by Cabot & Dennis (2011:2) are boosted by 1.5 dB to compensate for the relative gain at the positioning each side of the listener's head. In addition, each channel's power is summed in order to give an overall power rating for the full audio signal.

The K-weighting curve (fig. 2.7) shows that frequencies lower than 100 Hz undergo attenuation, between 100 Hz and 1000 Hz the frequency level is preserved and frequencies higher than 1kHz undergo amplification of 4 dB (Fleischhacker, 2014:6).

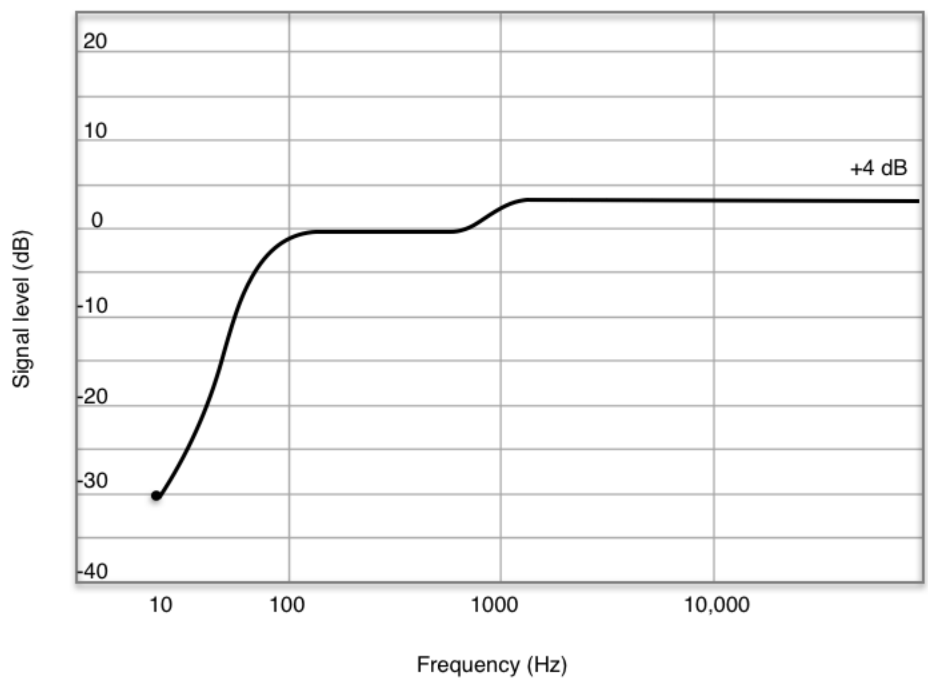


Figure 2.7: K-Weighting Filter (Fleischhacker, 2014:6)

Camerer (2010:8) and Cabot & Dennis (2011:2) highlight that the ITU-R BS.1770 algorithm was revised in 2011 resulting in an 'Integrator' extension (fig. 2.8).

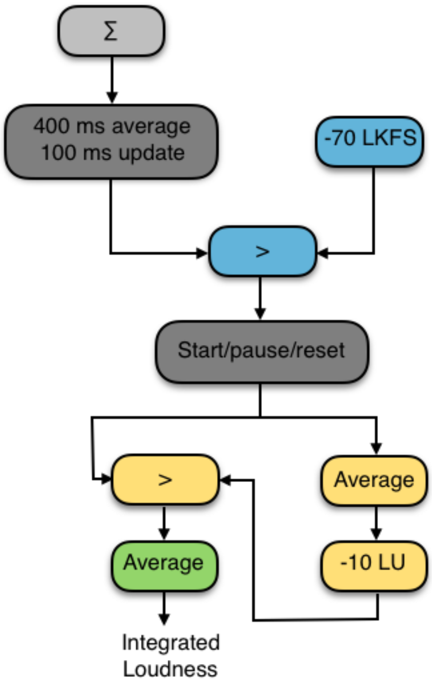


Figure 2.8: Revised ITU-R BS.1770 Extension (Cabot & Dennis, 2011:2)

The extension to the BS.1770 algorithm (fig. 2.8) works through averaging the power over a 400 ms window and is updated every 100 ms. Cabot & Dennis (2011:3) and Robjohns (2014:114) point out that this averaging works with a 75% overlap, resulting in the process undergoing further adjustment or weighting by means of a start/stop gating method. This allows for the analysis of a segment from the audio signal. In addition, an absolute gate of -70 LKFS is applied to ensure the elimination of fade outs and lead ins from the audio signal. The remainder of the extended algorithm, shown in yellow (Cabot & Dennis, 2011:3) focuses solely on the foreground audio and applies a two-step averaging procedure. This is done by averaging the 400 ms values over the entirety of the measured signal, with the result being reduced by 10 LU and used further as a gating threshold.

Lund (2013:1) proposes the BS.1770-3 to be implemented as the worldwide cornerstone of loudness normalisation. This is because the algorithm provides reliable discrimination between the foreground and background audio through it's measurement gating method. This algorithm works across all genres, platforms and audio formats regardless of whether the audio is linear or wide range.

The revised ITU-R BS.1770 algorithm has a multitude of uses within the music industry. MusicLoudnessAlliance (2012:3) points out that the solution of inconsistent loudness playback resides in the massive adoption of digital file based music. Due to all playback devices containing a computer chip, they can all analyse a sound file's average perceptual energy and thereby automatically control the output level. From an international broadcasting standpoint, the ITU-R BS.1770 revised algorithm is ideal for the prediction of subjective loudness.

The BS.1770 recommendation is therefore incorporated into both radio and TV broadcast to ensure overall loudness normalisation. Fleischhacker (2014:4) highlights that the ITU-R BS.1770 algorithm has already been implemented into the Association of Radio Industries and Businesses (ARIB) in Japan; OP-59 by Free-TV in Australia; Advanced Television Systems Committee (ATSC) in the US and EBU-R128 and Tech 3341 - 3344 in Europe. The notion of loudness normalisation is mandatory across all audio broadcasting industries to ensure that the listener is satisfied with the musical playback. Furthermore, the inconvenience of the constant adjustment of playback volume is adverted, as well as a reduction in the risk of hearing damage from loud musical changes and advert interjects.

The implementation of loudness normalisation into television shows the potential and control of the algorithm, supporting the progression into radio. Since the problems encountered with loud interjects exist both in radio and television, the addition of the algorithm into the latter provides a convincing argument for radio broadcasters to follow. In the case of this experiment, the results should yield the improved experience due to the support of loudness normalisation.

The following sections will focus on loudness normalisation algorithms, both built independently and based on the ITU-R BS.1770.

2.5.2 EBU-R128

The ITU and EBU are developing loudness standardisations to minimise current discrepancies in loudness within audio transmission both over radio and television. There are currently methods which can be used to standardise the output of musical loudness. Camerer (2010:1) recommends the use of EBU R128 as the defined method to measure loudness level for music, TV and films. However, as it currently stands, the R128 does not make sufficient provision for less than optimal listening environments. Less optimal environments include passenger cars, as many of them still have either a loudness button or loudness equalisation option within the audio settings (For example: Renault Clio 2006 Model, Jeep Renegade 2015 Model). Since loudness buttons exist in cars to begin with, it shows that listeners have had the choice to implement loudness normalisation. The algorithms however, provide an autonomous loudness normalisation that requires no knowledge of audio processing, or interaction from the user.

It is recognised that the EBU-R128 is a

> "defined method to measure the loudness level for news, sports, advertisements, drama, music, promotions, which helps professionals to create robust specifications for ingest to a multitude of platforms" (EBU, 2016:39).

Through a wider implementation of this recommendation, audio broadcasts will be standardised, thus creating a better listening experience for the audience.

The implementation of the R128 does however create a significant switch-over from the current broadcasting standard, but will positively impact both the organisation and economic parts of the broadcast industry (EBU, 2010:5). Whilst the implementation does require a significant change in broadcast, the EBU (2016:39) states that the R128 specification aims to make the measurement of audio compatible across the globe.

The EBU developed a recommendation built upon the ITU-R BS.1770 (Adriaensen, 2011:15), through the addition of three separate parameters: Loudness Range (LRA), True Peak Level (TPL) and Programme Loudness (Camerer, 2010:3).

The first parameter, the LRA, quantified in Loudness Units (LU), is used to measure loudness distribution over the entire audio track (EBU, 2016:40). According to the EBU (2016:18), the LRA is most affected by an individual's listening environment, age and what is considered to be their comfort zone. This means that using the LRA to justify the perfect playback level for loudness normalisation isn't adequate as the result will be different for each listener.

The second parameter, the TPL, is defined by Camerer (2010:4,5) as the highest value in a signal waveform, (either positive or negative) within a continual time period. The resulting value is often higher than that of a Quasi-Peak Programme Meter (QPPM) reading as the TPL value can only be detected by a meter compliant with the BS.1770 algorithm. This is because the BS.1770 compliant meters use oversampling to provide a better estimate of the TPL than the QPPM (EBU, 2016:42).

The third parameter developed on the BS.1770, is the Programme Loudness level. This refers to a singular value in LUFS (EBU, 2014:5) that works alongside the Target Level for the broadcast. In other words, the Programme Loudness level depicts the integrated loudness level over the entire programme and is compared to the broadcast Targeting level, which stands at -23 LUFS with a discrepancy of ± 0.5 LU (EBU, 2016:40).

The LRA and TPL of an audio track give the researcher an indication of how loud the track can be played prior clipping and distortion. The tracks used in this investigation have been analysed according to the guidelines of the TPL and LRA to ensure the correct control over the playback levels. The Programme Loudness level for this experiment was set to -23 LUFS in accordance with the EBU R128. More information about the audio processing of the audio tracks can be found in Chapter 3.

The EBU (2010:3) states that loudness inconsistencies between channels are the cause of most viewer and listener complaints. In an effort to reduce dissatisfaction amongst listeners, the EBU-R128 needs to be spread out across the music industry in a mass attempt to allow music producers to mix according to a worldwide standard. It is through this process that listeners will have less trouble with loudness discrepancies between musical tracks and radio channels.

Adriaensen (2011:11) points out that as part of the EBU initiative, the R128 will be available as a basis for audio processing software. The integration of the EBU initiative is evident in the MLoudnessAnalyzer (MeldaProduction, 2009) (fig. 2.9).

The software plugin utilises the EBU R128 standard whilst analysing the audio clip. The plugin allows the user to define the preset with the choice of EBU+9; EBU+18, EBU+27 and LUFS EBU R128. The use of analysis software allows for broadcasters and producers to see the quality of the audio tracks prior to transmission.

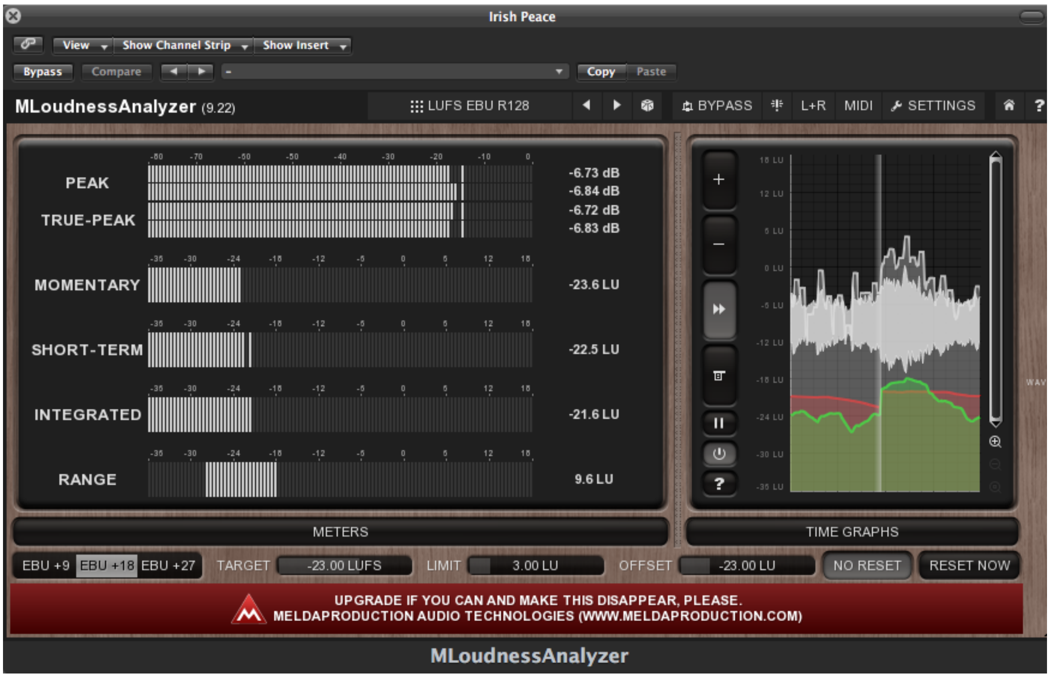


Figure 2.9: MLoudnessAnalyzer Measuring Loudness Parameters in Logic Pro (MeldaProduction, 2009)

For this investigation, the MLoudnessAnalyzer was utilised initially for analysis only, but was later discarded in favour of R128x software which allows for the processing of audio in accordance with the EBU R128. The MLoudnessAnalyzer provides an in-depth analysis of the loudness parameters characterising a chosen track, however does not freely provide the ability to export the chosen track in accordance with the EBU R128.

The EBU R128 was chosen as the loudness normalisation standard for this investigation as South Africa is moving toward a complete integration of the R128 over the next few years. Asikhule (2014) points out that during the African Loudness Summit of 2013, MultiChoice made the statement, accepting the integration of the EBU R128. MultiChoice announced that,

> "advertising content that complies to the EBU R128 loudness recommendation that is delivered digitally via LaserNet's Media Move service or via Adstream will be broadcasted by DSTV without any further audio processing" (Asikhule, 2014).

Asikhule (2013:2) highlights that the implementation of the R128 took place during August 2013, with the expectation that content producers will acknowledge the superior quality of higher fidelity programmes and begin to quickly cross over to the R128. Therefore, utilising the EBU R128 in this investigation will help provide further insight into the radio broadcast of high dynamic recordings.

In addition to the EBU R128, there are more independently operated loudness algorithms which are detailed below.

2.5.3 Replay Gain

According to Wolters & Riedmiller (2010:7) and Tagtaum (2016), Replay Gain is a non-proprietary loudness control algorithm developed in 2001 that is available in two versions: peak signal amplitude and gain adjustment.

Nygren (2009:10) highlights the difference between the two versions: peak signal amplitude aims to create loudness uniformity by calculating the gain correction of a single track and then applying it to the next track, whereas gain adjustment aims to calculate the gain correct value over an entire album. Wolters & Riedmiller (2010:7) highlights that peak signal amplitude is best suited for when individual tracks are played in a mix from a variety of albums, such as over radio broadcast, whereas the gain adjustment version is best suited for when all tracks of a singular album are played consecutively, such as domestic listening.

There is a marginal difference in filtering between the BS.1770 and Replay Gain as Wolters & Riedmiller (2010:8) points out that the BS.1770 essentially applies a high pass filter, whereas Replay Gain uses a band-pass filter. Nygren (2009:10) points out that this band-pass filter incorporated by Replay Gain looks similar to an inverted approximation of the Fletcher-Munson curves.

The Replay Gain versions are more widely accessible as seen by an application called Beatunes (Tagtaum, 2016). BeaTunes allow users to analyse their musical library and apply Replay Gain to their tracks. Since Replay has both versions for track-track and album-album, it is understood that Replay Gain is an ideal alternative to the BS.1770 for domestic listening. Wolters & Riedmiller (2010:7) suggest that the adoption of Replay Gain as a syntax for loudness measurement and control is a great idea as it aids the spread of loudness control within the industry. The most ideal integration would however be through Replay Gain where users can match the semantics of Replay Gain with the BS.1770 (Wolters & Riedmiller, 2010:7). This will allow a uniform Target Level loudness output across the board. The MusicLoudnessAlliance (2012:3) states that adopting the ITU-R BS.1770-2 into Replay Gain would be most advantageous as it means that loudness normalisation would then be based upon a single international standard.

Whilst Replay Gain is commercially available both for the implementation by radio stations and for domestic listening, Apple's SoundCheck is available to any user with an iOS device or Mac computer.

2.5.4 Apple Sound Check

Apple's Sound Check function has been available to all Apple users since iTunes 3 was introduced during 2002 (Robjohns, 2014:1). The Sound Check function is accessible from within the playback settings in iTunes. When activated, Sound Check scans the individual's music library and stores the loudness values of the tracks in the audio file's metadata. From personal experience the analysis of one's library can take several minutes depending on the size of the music library. This process is bypassed if the individuals purchase their musical tracks directly from the Apple Store (Apple, 2012:5).

Similarly to Replay Gain, the objective of Sound Check is to stabilise the loudness spikes between audio tracks across all iOS/OS devices. Sound Check is however limited to specific audio formats as according to Apple (2016), "Sound Check works with .mp3, AAC, .wave and .aiff file types" and therefore will not work if the audio track is of an unsupported format. From a competitive perspective, this could be seen as a disadvantage because the ITU-R BS.1770 recommendation as well as the Replay Gain algorithm both work across all systems and all audio formats.

Lund (2013:4) states that Apple's Sound Check has a favourable feature which is able to normalise loudness variations across both old and new tracks without having been based on the ITU-R BS.1770 algorithm. The Sound Check is instead, solely developed within Apple and the specifications unknown to the general public.

Whilst the Target Level of Sound Check is unknown to the public, Lund (2013:4) and Robjohns (2014:1) both state that the median Target Level of Sound Check on a BS.1770-3 scale gives a reading of approximately -16.2 LUFS. This is a favourable outcome for Apple Sound Check since it is not based on the BS.1770 algorithm. Robjohns (2014:1) does however highlight that the playback volume when utilising Sound Check is inadequate. Furthermore, in order for the loudness normalisation to be most effective, the target level within the algorithm needs to be low enough to accommodate the highest peaks likely to be encountered within the music library. Since the target level is at -16 LUFS, the output volume for compressed audio tracks is considerably softer than the playback level when Apple Sound Check is switched off. This deviation between playback levels can, according to Robjohns (2014:1) be up to 12 dB softer, which is a considerable attenuation level.

From personal experience, through learning how to use the Sound Check feature within iTunes, the output level of a Macbook Pro gives a significantly softer output than if the tracks were played without the Sound Check function. This discrepancy could be reduced drastically through the implementation of make up gain prior to playback. It should be noted that especially within the European Region, this discrepancy can cause a marked inconvenience to iOS users as the musical playback on iPhones can be limited further to the European Union playback level (fig. 2.10).

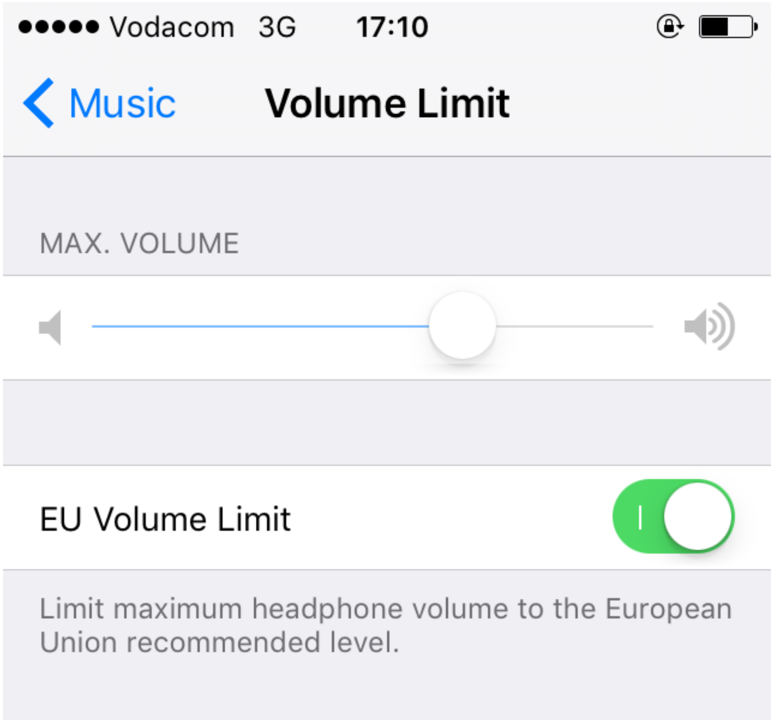


Figure 2.10: European Union Volume Limit, iPhone 5S (Personal Phone)

The volume limit option is found in the Music Volume Limit settings on the most recent iPod and iPhones. In this instance, the volume limit is set to enabled on an iPhone 5S. Therefore, if the Apple user has both the Sound Check and EU Volume Limit active, the output volume level is greatly reduced and for some audio files, the output volume can then become too low. Lund (2013:4) supports this notion of audio files not being loud enough as he was listening to a BBC radio podcast on an aeroplane journey with the Sound Check enabled, which resulted in the podcast not being heard loud enough to be listened to comfortably over the background noise generated by the aeroplane's engines.

2.5.5 MasterCheck

In contrast to the loudness algorithms discussed above, MasterCheck is a loudness meter designed by NugenAudio to facilitate high-quality audio within the loudness normalisation paradigm. NugenAudio (2014:3) highlights that MasterCheck incorporates the use of dynamic monitoring and an ITU compliant true peak inter-sample meter. This aids the sound engineer with producing a track compliant with international loudness standards. Current consumer audio codecs such as MP3, Ogg and AAC introduce clipping into the audio signal which hinders the highest quality outcome for the artist. Therefore, by the implementation of the True Peak Levels as part of the ITU-R BS.1770 into MasterCheck, the artist is provided with a more accurate signal reading than traditional peak metering systems, leading to more accuracy prior to the audio transmission (NugenAudio, 2014:3).

Whilst this implementation of a loudness standardisation works best at the transmission stage, NugenAudio (2014:3) states that it does present a disadvantage for the producers. At playback, the hot master effectively becomes diminished, thereby eliminating the advantage and attempt to be louder than audio competitors. Furthermore, MasterCheck works in contrast to Replay Gain and Apple Sound Check as the function happens at the mastering stage prior to the audio production over radio broadcasting. As discussed above Replay Gain works by normalising peak signal amplitudes as well as adjusting the gain levels and Apple Sound Check normalises audio codecs to a set loudness standard, whereas MasterCheck solely provides a visual metering system for the producer to adjust their mix at the mastering stage.

2.5.6 Algorithms in Conclusion

It is important to detail the independently operated loudness normalisation algorithms as they provide an accessible listening function for the majority of listeners. In the case of Apple Sound Check, the function is available to any individual in possession of an OS/iOS device, ReplayGain can be implemented into radio stations or used by audiophiles and MasterCheck can be utilised by music producers.

As it stands, only the large broadcasting services, specifically Multi-choice/DSTV in South Africa show implementation of loudness normalisation, whilst smaller local stations still make reference to Peak Programme Meter (PPM) values (Loots, 2016:19). Since the local broadcasting industry has not yet fully incorporated loudness normalisation, there is more emphasis on the individual listener to control their loudness playback. Individual listeners may incorporate loudness normalisation to their music libraries through the implementation of Replay Gain or Apple SoundCheck. Furthermore, the use of MasterCheck could be beneficial to independent producers or for individual's producing their own musical material. This way the industry as a whole can benefit with each sector gradually becoming more familiar with software aiding the use of broadcast normalisation standards.

Through the vast implementation of both independently developed and international standardised algorithms the audio spikes across the music and broadcast industry can be reduced at a quicker rate.

2.6 Loudness Descriptors

With the implementation of loudness algorithms and the introduction of metering systems, the use of loudness descriptors provide another method of understanding the loudness of audio tracks. This section will focus on loudness descriptors and how they visually represent loudness.

Loudness descriptors are defined by Skovenborg & Lund (2009:2) as, "[...] key numbers to summarise loudness properties of an audio segment, broadcast programme or music track". These numbers are calculated through different loudness visualisation programmes which allow the engineer to appropriately monitor the loudness output of any given programme or audio track.

Prior to the implementation of loudness descriptors, the broadcasting stations would utilise both analogue and digital meters in an attempt to understand the loudness of a given track. Due to the limitation of both the current analogue and digital meters, loudness descriptors, in the form of visualisation meters, have been developed to give a well rounded representation of a given audio segment's loudness.

2.6.1 Digital Meters

The analogue VU meter is outdated in comparison with other metering systems because it displays an absolute value over a slow time constant. Skovenborg & Nielsen (2007:2) argues that due to the underlying algorithm, the VU meter is inadequate to deliver a reading suitable for describing the loudness of a given sound. The replacement digital meters, called Peak Programme Meters (PPM) are capable of displaying a peak reading, which according to Orban (2000:22) is better suited to monitor operating levels where a small amount of clipping is present. Skovenborg & Nielsen (2007:2) highlights that PPMs come in two forms, one with an instantaneous response and the other with a short (few milliseconds) response time to rising levels.

Both the VU and PPM systems do however play a role in recommendations pertaining to audio broadcast. For example, the VU metering systems are utilised mostly in America and Australia, whereas Europe utilises PPM systems for radio broadcast. Furthermore, Spikofski & Klar (2004:3,6) state that a Quasi-Peak Programme Meter (QPPM) is utilised in the International Electrotechnical Commission (IEC) recommendations as a sub category of PPMs. This is due to the ability of QPPMs to ignore shorter duration variations in the signal.

A visual comparison between the uses of QPPMs and VU metering systems displaying the necessary headroom requirements for an ideal reading is illustrated in Figure 2.11. It can be seen that QPPM systems require only 9 dB of headroom, whereas VU systems require up to 18 dB, significantly more than QPPM systems. The headroom allocation for each system ensures the attack time of the meter which indicates that QPPM systems will respond at a much quicker rate than VU meters. Spikofski & Klar (2004:3,4) highlight that the IEC QPPM system used by the BBC takes above 10 ms to reach the 80% tag line, meaning that after 10 ms, the system is operating at 80% of digital full scale.

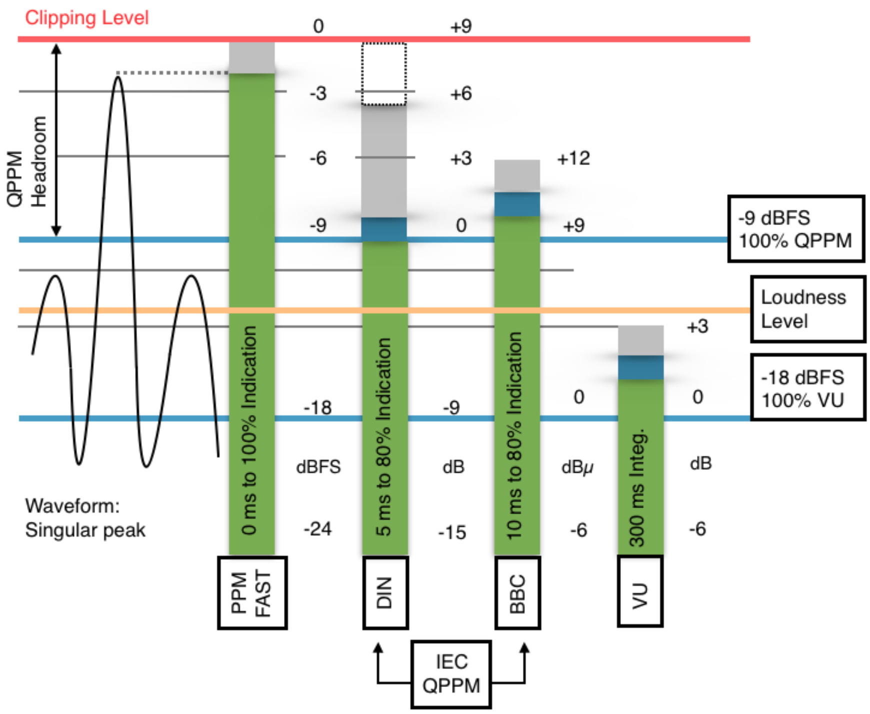


Figure 2.11: Characteristics of Broadcast Metering Systems (Spikofski & Klar, 2004:3)

Orban (2000:23) states that each broadcasting station has a unique set of guidelines that govern the use of PPMs to give the subjective and artistic sound that is desired. This is evident in the example (fig. 2.11) used by Spikofski & Klar (2004:3), illustrating that the attack time of PPM systems used by German broadcasters reach a higher tag line percentage than the BBC within the same 10 ms time period. The outcome therefore shows that both German and British broadcasters use different subcategories of PPMs to achieve the desired outcome when mixing audio for their audiences.

Whilst both VU and PPM systems are utilised within the broadcasting industry, Skovenborg & Nielsen (2007:3) highlights that loudness is a perceptual quality of sound and can therefore be modelled using specific psychoacoustic algorithms giving an objective calibre. These psychoacoustic algorithms provide the producer with a different perspective, presenting loudness characteristics unavailable through the use of PPM systems. Therefore, using objective visualisation in addition to PPM systems allows for a comprehensive understanding of the audio qualities.

The next section will highlight the different forms of loudness visualisation demonstrated through loudness descriptors.

2.6.2 Loudness Visualisation

The developments in loudness descriptors contribute to the understanding of loudness properties of an audio segment.

The first descriptor, Center of Gravity (CoG), refers to the measurement of the overall loudness of a given audio segment. Skovenborg & Lund (2009:3) points out that the CoG works as an integrating loudness measurement, similar to the L_{eq} measurement used within the ITU-R BS.1770.

The COG measurement incorporates an adaptive gate that is stricter on the quiet segments within the audio track. This essentially means that the CoG measurement ignores the quietest, potentially inaudible segments of the audio track. The inaudible segments are purposefully ignored as their implementation would generate a biased result by the CoG measurement (Skovenborg & Lund, 2008:2). The use of an adaptive gate implements a relative gating threshold that, according to Skovenborg & Lund (2008:2,3) prevents the possibility of a fixed threshold cutting out the softer sections of music across all genres.

With the use of a fixed threshold at a high level across a variety of genres, a problem arises because the softer sections across the genres vary in loudness. This means that a fixed threshold might work fine for one genre, but cut out parts of the softer sections in another. Skovenborg & Lund (2009:3) points out the solution, through the implementation of an adaptive gate set with a relative threshold set at -20 dB. The -20 dB relative gating threshold prevents the threshold getting stuck at higher amplitudes as well as rendering good results across a variety of musical genres.

The second descriptor, Consistency, is defined by Lund (2009:34) as loudness changes within a musical track. The Consistency descriptor is also measured in LU as opposed to the CoG unit of LKFS. Lund (2009:34,35) states that consistency gives an objective measurement of a musical track's loudness range (LRA), whereby the LRA is equal twice the Consistency. The Consistency scale reveals that 0 defines the top of the scale, indicating an input signal or tone. Therefore, in order for a track to have a constant loudness output, a fader gain ride (fig. 2.12) is implemented over a 10 dB range (+5 dB to -5 dB).

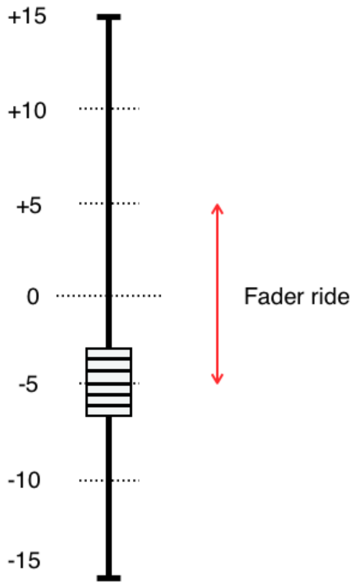


Figure 2.12: Constant Loudness via Fader Gain Ride (Lund, 2009:36)

The fader gain ride regulates the musical track's loudness variations. Common examples of Consistency readings from commercials lie within the range of -2 LU to -0.5 LU (Lund, 2009:39), whereas 80s rock/pop songs lie between -7.8 LU (Pink Floyd - Shine on Your Crazy Diamond) to -1.8 LU (Ry Cooder - Don't mess up a good thing) (Skovenborg & Lund, 2008:8). The discrepancy of the loudness (in LU) shows the dynamic range quality of each production as well as the necessity for the loudness normalisation algorithms. Through the comparison of Consistency readings, the musical track with the most appropriate dynamic range can be selected for the present experiment.

Using the same examples above (Pink Floyd and Ry Cooder), the CoG values for each track are -15.9 LKFS and -24.2 LKFS respectively Skovenborg & Lund (2008:8). With a variation in both the CoG and Consistency values, creating a standardised loudness output for the listener becomes more of a challenge. However, keeping the loudness output to a manageable range is an attainable goal with the implementation of loudness normalisation algorithms.

The final descriptor gives the most appealing visual description through displaying the Short Term and Long Term loudness of an audio segment on a radar view. Lund (2006:61) describes that the radar view objective is to, "produce an accurate and robust estimate of the perceived loudness of sound segments consisting of both speech and/or music". Furthermore, Lund (2006:61) points out that with development, the concept will operate with realtime short and long term loudness, giving a more accurate and detailed description of the perceived loudness of an audio segment. The full concept can be seen below (fig. 2.13).

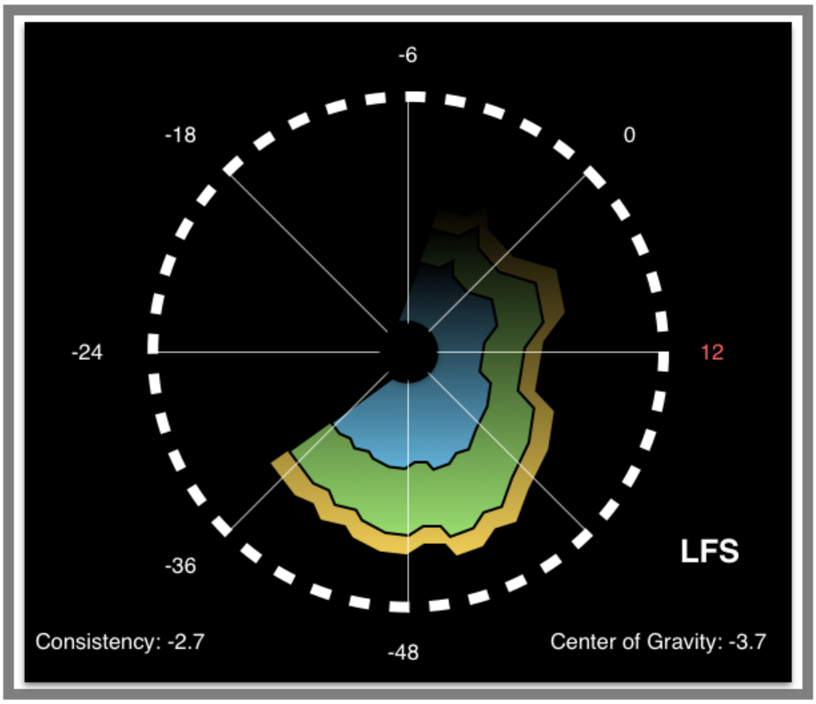


Figure 2.13: Loudness Radar Displays Long and Short Term Loudness (Lund, 2009:45)

The graph displays a circular wheel along the outside with a coloured spread that operates in a clockwise direction across the middle. Skovenborg & Nielsen (2007:4) describes that short term loudness, displayed as the colours, allows the operator to deduce whether the audio segment is at the ideal loudness, being neither too soft nor too loud. Moreover, the scale running along the top is linear and is similar to the LU scaling within the ITU recommendation. This linear scale runs over a larger range than that of the ITU recommendation, from -20 LU to +15 LU. In the example above however, the scale is shifted towards the lower region.

The Short Term loudness, as specified by Skovenborg & Nielsen (2007:4) represents both the size of the arc as well as the position of the curved bar (in the above example, the bar points to about -30 LU). Should the producer or operator want to compare multiple tracks at the same time, these properties of Short Term loudness allow for a more in-depth comparison. The Long Term loudness works through the movement of the arc's bar following a clockwork motion. The most recent loudness reading marks the brightest colours, with the older loudness reading fading to black at the arc's tail. Furthermore, the distance the arc moves away from the centre of the display, the louder the sound is perceived at that given point in time. Skovenborg & Nielsen (2007:4,5) states that the time period in which this descriptor operates can be adjusted to scan both the past minutes as well as the past hours through zooming in and out respectively.

The descriptors discussed above work in unison with the loudness algorithms and help display an objective measurement from a listener's perception of an audio track. Through incorporating a loudness concept such as this, the possibility to create the best volume output that suits the majority of listeners becomes a plausible goal.

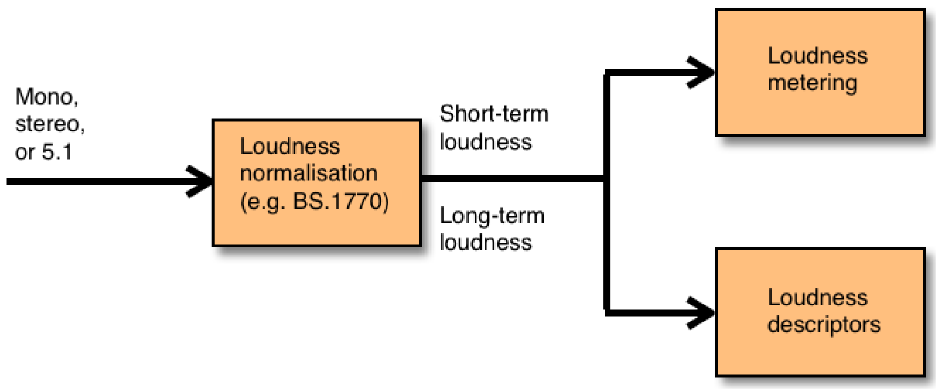


Figure 2.14: Loudness Chain: Measurement, Metering and Descriptors (Skovenborg & Lund, 2008:3)

Figure 2.14 above, shows the processing chain of an audio file, measuring the loudness output of single/multiple audio tracks or unifying the loudness across all audio sources.

The loudness visualisation wheel for displaying both short and long term loudness is still in the prototype stage as the meter does not align the levels without the use of a human operator (Skovenborg & Nielsen, 2007:1). Therefore, for the description of an audio track's loudness, the use of the CoG integrated loudness as well as the Consistency, gives the operator a better loudness description. In order to get the best description of the loudness within an audio track, both the descriptors are used alongside the visualisation meter (fig. 2.14).

2.6.3 Loudness Integration

The above sections described the subjectivity of loudness and how it can be portrayed through an objective measurement. The use of an objective measurement is available to the consumer, should they choose to implement loudness normalisation algorithms to control their musical collections. The quality of audio formatting plays a large role in the dynamic value or musical tracks. In order for these to be shown, a Dynamic Range Tester needs to be used to calculate the hotness of the track's dynamic material.

The hotness scale of an audio track or album can be viewed on the DynamicRangeDatabase (2014). If the album or track does not appear on the site, the Dynamic Range Database gives the user an option to download an offline version to test the album or tracks themselves. The offline meter determines the Dynamic Rating of the track or album, which can be used together with the database's dynamic scale to show the album or track's dynamic range.

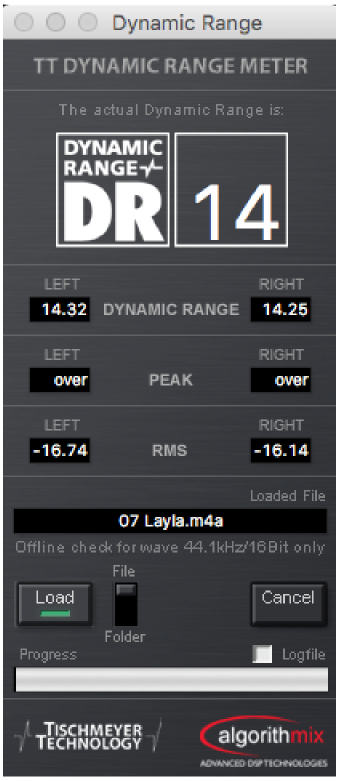


Figure 2.15: Offline Dynamic Range Meter (DynamicRangeDatabase, 2014:1)

The offline meter (fig. 2.15) was used to analyse Eric Clapton's Layla, which can be compared to the scale presented on the Dynamic Range Database website (fig. 2.16).

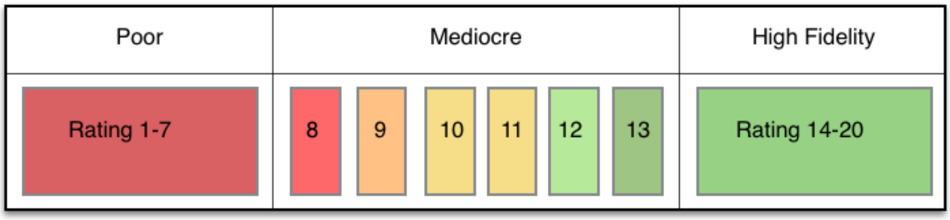


Figure 2.16: Dynamic Range Scale (DynamicRangeDatabase, 2014:1)

It can be seen that the offline Dynamic Range Meter (fig. 2.15) gives a Dynamic Reading (DR) for each of the tracks uploaded by the user. In the example above, the Layla track has outputted a DR of 14, which, according to the Scale (fig. 2.16) shows a good dynamic range for the track. As mentioned above the offline meter can also scan albums, which gives a DR for each track including an average DR for the album.

The Dynamic Range Meter is useful because it shows how the audio format affects the broadcasting capability of each track or album. Should the track have a good dynamic range, as most classical tracks do (Gymnopedie No.1 has a DR14), the broadcaster knows that the appropriate environment would include one with a lower noise floor. This would allow for the classical track to be played and enjoyed at an suitable volume. The tracks could be played in an environment with a higher noise floor, provided the musical track undergoes the correct amount of loudness normalisation.

2.6.4 Loudness Descriptors in Conclusion

The Dynamic Range Database (fig. 2.16), coupled with the Offline DR Meter (fig. 2.15), were used in this investigation to find the tracks best suited for the experiment. The display of an overall DR value for both single tracks and full albums helped the researcher determine which available tracks had the highest dynamic range.

The loudness visualisation wheel (fig. 2.13) is not yet available for commercial use. Therefore, this method of loudness visualisation will not utilised in this study. However, when the development of the objective visualisation becomes available, the utilisation would be most beneficial for displaying the loudness characteristics of musical tracks.

The present research utilises the L_{eq} settings from the DSP Mobile Analyzer, displaying a graph of the peak amplitude and frequencies each participant is exposed to. In order to establish a musical track's dynamic range, the Dynamic Range Database was used. This is due to the lack of access to the Consistency loudness descriptor hardware and software.

Chapter 3

Research Methodology

3.1 Research Design Overview

This experiment took the form of a qualitative study recording the responses of the participants. This was done through the observation and recording of the volume adjustment responses made by subjects listening to musical tracks from the driver's seat of a car. The data collected was then analysed quantitatively. Thirty nine individuals participated in a listening task as part of the experimental procedure. Each listener was seated in a passenger vehicle, surrounded by active speakers playing pink noise, thus simulating the background noise level of a city-centre. The subject then had to listen to two musical tracks from two different genres, one popular rock and the other chamber music. Each of the tracks were of highest audio quality and processed in accordance with the EBU R128 at -23 LUFS. This gave both tracks the same average dBA output.

Each individual was asked to fill in a short questionnaire to help build up a demographic of the participants. This was done under the supervision of the researcher prior to the experiment commencing.

The details of the experiment were explained to each subject in terms of their preference in loudness for each track they had to listen to, whilst seated in a car. Prior to the beginning of each recording, the researcher explained the experimental procedure ensuring there were no surprises for the subjects. The subjects were given the opportunity to ask questions about anything they were uncertain about prior to the start of the experiment.

3.2 Hypothesis

For the purpose of this research the following hypothesis is stated:

> Whilst exposed to elevated levels of background noise, subjects will increase the loudness level of each audio track to a level high enough to be enjoyed over the background noise.

Based on the hypothesis, the following aims were set for this experimental procedure:

- To determine the preferred loudness playback level for passenger cars drivers whilst listening to two audio tracks against background noise
- To determine the optimal loudness range for radio transmission in a passenger car

The individual volume selections of each participant will be used to generate a statistical p-value, linking the two musical tracks. Determining the loudness playback value, coupled with a statistical p-value, may help in monitoring the loudness normalisation over radio transmission.

3.3 Subject Selection

3.3.1 Criteria for Subject Selection

The selection of subjects for this experiment required no fixed demographic criteria other than that hearing should be within normal limits, thus not experiencing any hearing difficulties. The subjects were randomly selected by approaching students on their way to class, as well as contacting students in residence directly and inviting them to participant in the research. Individuals contacted directly were asked to bring colleagues or friends to expand the subject group. Appendix C displays a breakdown of the experimental details. A sequential breakdown of the experimental procedure was presented to the subjects upon contact and displayed on the car window during the experiment.

3.3.2 Subject Selection Procedures

A self administered questionnaire (in the presence of the researcher) was used to determine individual demographic information, as well as to determine the individual's knowledge and insight into loudness normalisation algorithms. The questionnaire was carried out by each participant under the supervision of the researcher, prior to the playing of musical tracks. The full questionnaire is presented in Appendix A, with the motivation for the selection of each question shown in

Appendix B. Each participant was made aware that answering all questions in the questionnaire were not a pre-requisite for continued participation in the research.

It was felt that the majority of participants would enjoy listening to music in their vehicle as well as agree with the researcher's thoughts of adverts and interjects broadcasted over the radio being too loud. Furthermore, the constant adjustment of musical track loudness is felt to be a pressing concern by the researcher, with hopes that the participants would share his concern.

The desired listening level was to be determined with each volume adjustment by the participants. Riedmiller & Robinson (2003:3) state that the comfort zone defines a loudness range whereby the listener accepts the volume discrepancies between audio tracks. The implementation of loudness standards provides uniform audio loudness levels across the broadcasting industry for the comfort of the listeners. Therefore, this information was shared with the subjects to provide motivation for their participation.

3.3.3 Sample Size

A sample of 40 was deemed suitable in terms of the scope of the study and the amount of people which could be accommodated during the completion of the experiment. Of the proposed 40 sample size, 39 individuals participated in the research.

3.4 Experimental Layout

The experimental procedure and design, as well as the stages of execution of the experiment are presented below.

3.4.1 Audio Track Development

The tracks were: Another Brick in the Wall Pt.2 by Pink Floyd[1] off The Wall album and the Violin Concerto No.3 by Camille Saint-Saens[2]. These two tracks were chosen to represent the two genre extremes of popular rock and classical music. The initial tracks selected included:

- Pink Floyd: Another Brick in the Wall Pt.2
- Eric Clapton: Layla, Unplugged
- Mark Salona: Rafael
- Wolfgang Amadeus Mozart: Piano Sonata No.11

[1] According to Urick (2016), The Wall album was released in 1979 and was deemed one of the more creative albums in rock music

[2] MusOpen (nd) highlights that Saint-Saens composed his final Violin Concerto with a sense of subtle impressionism in 1880.

Camille Saint-Saens was chosen over Rafael and the Piano Sonata because it was previously recorded by Sun Studios as part of Stellenbosch University in accordance with the -23 LUFS. Another Brick in the Wall Pt.2 contained a good variety of both vocals and instrumental music which outweighed the solo voice and accompaniment of Layla.

Both the Violin Concerto No.3 and Another Brick in the Wall Pt.2 had 30 second segments selected as it was felt that a 30 second window would be long enough for a listener to make their loudness decision. The time window was tested briefly by the researcher with 1 minute and 45 second segments respectively, but was felt to be too long as the loudness indication was almost instantaneous.

3.4.1.1 Audio Processing of Tracks

The musical tracks were played to the participants as the artist/composer intended. This means that the tracks did not undergo any dynamic compression. The only processing Pink Floyd and Saint-Saens underwent was in accordance with the EBU R128, to ensure uniformity of the output levels to -23 LUFS.

This procedure was carried out in ProTools using the WLM plugin (WavesAudio, 2016) with the newly processed files analysed further with the R128x (Github, 2016), ensuring the Integrated loudness outputs for both tracks was set to -23 LUFS. Once the tracks had been aligned in accordance with the EBU R128, both tracks were checked again using the R128x by Github (2016). The difference in musical dynamic range is explained in LU, whereby the Pink Floyd has 2 LU, compared to the 14 LU of the Saint-Saens. This means that the Pink Floyd track was recorded and produced with more dynamic compression, whereas the Saint-Saens has no added destructive dynamic compression.

The method of audio processing through the WLM plugin was chosen over the Auphonic batch processor (Auphonic, 2016) as a more trusted source within ProTools was deemed the better option. Furthermore, the MLoudnessAnalyzer (MeldaProduction, 2009) for Logic Pro was discarded in favour of the R128x analysis software. This is because the MLoudnessAnalyzer presented limitations due to it being a trial version.

3.4.1.2 Background Noise

In addition to the two selected musical tracks, an hour long wave file of Pink Noise was generated using Tone Generator software (NCH, nd). The duration of the Pink Noise track was generated long enough to ensure that the track could play continuously without abruptly stopping midway through a subject's participation.

3.4.2 Research Set-up

This section will detail the experimental setup and the full procedure, from subject participation through to the conversion of linear data into dBA readings.

3.4.2.1 Research Equipment

The software preparation prior to the undertaking of the experiment, involved the choice of analysis software and microphone. The analysis was to be done using an iPad 2.0 equipped with a Thomann Mic-Wi436 (Thomann, 2016) and the Analyzer software from DSP Mobile (DSP-Mobile, 2012) installed. This analysis software was chosen over the initial Sound Meter by Faber Acoustical (FaberAcoustical, 2016) due to the extension support for the Mic Wi436, as well as a visual display graph instead of a simple decibel meter. The iPad, coupled with the Mic-Wi436 and DSP Mobile Analyzer were utilised during the post-experimental stage to convert the linear scale of the Renault's radio volume control to a decibel scale.

The Apple iOS 6 update from 2012 allows developers to bypass the high-pass filter limitation of previous iPhones. This enables Apple iPhones to connect with external microphones such as the Mic Wi436 which comply with the IEC 61672 Class 2 sound level meter standard (Kardous & Shaw, 2015:12).

Finally, the researcher tested both the musical tracks through Audacity and through an auxiliary cable into the car's stereo system ensuring that the tracks played without any disruptions. The microphone and DSP Mobile Analyzer software were also tested in detail to ensure accurate calibration between the sensitivity of the microphone and the DSP Mobile Analyzer software.

The full list of equipment used during the experiment is listed as follows:

- Radio Shack Digital Sound Level Meter
- MacBookPro with Audacity
- MacBookPro Charger
- iPhone to play Pink Noise
- iPad 2.0 with the DSP Mobile Analyzer installed
- Two Yamaha MSR100 Loudspeakers
- XLR Cables
- DI Box
- Jack to 1/4 inch Jack for iPhone to DI
- Multi-plug
- Extension Power Cable
- Renault Clio 2006 Model with Auxiliary port

The vehicle and speakers were set up in as shown in Appendix E, with each of the speakers placed on the ground next to the front wheel and front doors at a 45 degree angle. The speakers would generate pink noise to give the simulation of the the car being in city-centre traffic. The windows of the front seats on both the passenger and driver's sides were cracked open for air and the car engine was switched off during the time of the recording. This would allow for the right amount of pink noise to bleed into the vehicle. The Radio Shack (RadioShack, 2011:1) digital sound level meter was used alongside the pink noise from the speakers to calibrate the interior noise level. This was set to 60 dBA, measured with a slow response time. This gave the reference level within the vehicle prior to any music tracks being played.

3.4.2.2 Research Venue

The experiment was setup behind the Endler Hall of the Conservatorium at Stellenbosch University. This area was large enough to accommodate the car as well as being easily accessible for the participants. Moreover, the location was secluded from the main street, allowing for the background noise variable to better controlled. The vehicle was parked close enough to the Conservatorium's backdoor allowing for easy access to external power to supply the equipment.

3.4.2.3 Experimental Set-up

The experimental set-up is displayed in Appendix E. The speakers were connected to one another as well as to the multi-adapter for power. The speaker next to the passenger door was connected to the DI box and onward to the iPhone as this would provide the source of the Pink Noise. The researcher was seated in the passenger seat of the car, holding the iPhone and the necessary cables were running through a crack in the car window to power the MacBookPro on the researcher's lap. The MacBookPro was unplugged from the power source at the start of each participation as the power source caused interference with the car's stereo system.

3.4.2.4 Questionnaire Design and Completion

Prior to subjects participating in the experimental procedure, they were asked personally by the researcher if there were any questions as to what was expected of them or what was to be presented. In this way, each participant knew what to expect during the experimental phase, allowing for their comfortable input and relaxed loudness adjustment.

For each participant, the researcher would invite them to be seated in the driver's seat before prompting them with the Declaration (Appendix D) and Questionnaire (Appendix A). The declaration would be read through with the participant to ensure they were satisfied with what was expected, as well as to clarify any questions the participant had regarding the experiment. Furthermore, the questionnaire was filled out under the supervision of the researcher to ensure that each question was understood.

The questionnaire was designed to build a demographic of the subjects partaking in this study. The questionnaire can be seen in Appendix A, with a full motivation for the choice of each question in Appendix B. From a psychological point, questions 2 and 3 focussed on why some individuals prefer music louder or softer based on their exposure to music. This issue was briefly addressed, but the role of psychological attributes to loudness preferences was beyond the scope of this study.

After the declaration was signed and the questionnaire completed, the pink noise was turned on and the participant listened to each track. The Pink Floyd track was played to each participant first, followed by the Saint-Saens. Once participants verbally indicated that they were happy with their loudness selection, the procedure was terminated. The subject was then thanked for his/her participation.

It was decided that a pilot study was not necessary prior to the undertaking of the full experiment. The two musical tracks were pre-selected from laboratory tests whilst the rest of the research design is considered to be easily understood.

3.4.2.5 Data Recording

Once all subjects had participated, their selections were still in a linear format from the volume control on the car's dashboard. For example: Another Brick in the Wall Pt.2 played at volume number 25. In this linear format, the results show a variation across loudness levels, but cannot be used for comparison and therefore required conversion into decibels. This conversion procedure included the researcher seated in the driver's seat listening to each 30 second track and analysing the decibel peaks generated using the Mic Wi436 on the DSP Mobile Analyzer.

Even though a few participants had a cross over loudness preference, the linear value was run only once. This means that the linear values were run from 20 - 32 for the Pink Floyd track and from 17 - 32 for the Camille Saint-Saens track. The peak value for each 30 second segment was recorded as well as a CSV file generated for the L_{eq} over the entire segment of each reading. The DSP Mobile Analyzer was set with a slow response time, L_{eq} and 1/3 octave band for each generated CSV file. This data will be presented in the next chapter.

3.4.3 Ethical Considerations

Prior to conducting the experiment, ethical clearance was obtained from the University's research ethics committee.

The investigation incorporated the participation of human subjects for the evaluation of perceived changes in loudness between popular and classical music tracks. Each subject was presented with two forms outlining the purpose of the research as displayed in Appendix C, as well as the written consent for their participation as shown in Appendix D. Upon undertaking the experiment, the participants were asked to complete a questionnaire under the supervision of the researcher. Anonymity of the participants was ensured through the use of participant numbers instead of their names.

3.4.4 Budget

The researcher had access to literature, sound equipment and hardware needed for the purpose of this research. The software used in the experimental phase was bought by the researcher, amounting to R345 and was paid as a personal expense. The investigation was conducted using sound equipment from the Stellenbosch University's Studio, at an outdoor location large enough to work with a passenger car. The researcher accepted the responsibility of costs associated with the postgraduate study in his personal capacity.

Chapter 4

Results and Analysis

In this chapter the results obtained from the experimental procedure will be visually presented.

4.1 Questionnaire Results

The following is a full breakdown of the Questionnaire results. Table 4.1 shows the questions where participants could select from more than one option.

Table 4.1: Questionnaire Results

No.	Question Summary	Percentage Distribution	
1	Participant age	74% between 20 and 25	Average age 23
2	Vehicle Music Exposure	82% always listen	18% less often
	Music in Bars	56% occasionally attend	44% more often
	Live Performances	79% occasionally attend	21% more often
3	Daily Music Exposure	39% less than 2hrs	28% between 3 - 4hrs
		18% 4-6hrs	15% greater than 6hrs
8	Normalisation method	23% recognise	77% did not recognise
9	Popular audio formats	79% use MP3	38% lossless audio

Question 1: asked the participants to state their age, giving a demographic overview of the subject group. The results show a large age range from 19 to 38. From Table 4.1, 74% of the participants were aged between 20 and 25, with the average participant age being 23 years old.

Question 2 & 3: pertain to the noise related hobbies and exposure of the participants. This builds a demographic, but does not aid the measurement in any way. It was found that 82% of the subject group always listen to music,

with the remaining 18% occasionally listen to music in their vehicles. Furthermore, 56% of the subjects occasionally attended bars or clubs, whereas 79% attended live concerts or festivals as often as possible.
The results also show that 38% of the subjects are exposed to less than 2 hrs of loud music per day, with only 15% actively involved with loud music more than 6 hrs a day. It can be deduced that the overall noise exposure for this sample of participants is relatively low, which may attribute to the psychological loudness decision making when listening to the musical tracks.

Question 8: asked whether any of the participants were aware of loudness normalisation methods that are readily available. The results show that 10% recognised and have utilised Apple SoundCheck, 13% were aware of and had used the Loudness Button in vehicles, whilst the majority (77%) had no knowledge of loudness normalisation methods. This question was important as it ties in with overall implementation of loudness normalisation algorithms across the music industry. This is discussed further in the next Chapter.

Question 9: asked the participants to select which of the audio formats they used the most. The participants were allowed to select more than one option. It can be seen that the majority of the responses (79%) used MP3, whilst collectively the responses for the lossless audio format totalled 38%.

Table 4.2 below displays the results from the questionnaire that required a yes or no answer from the participants.

Table 4.2: Questionnaire Yes/No Results

No.	Question Summary	Answered Yes	Answered No
4	Music as a distraction	13%	87%
5	Radio interjects too loud	49%	51%
6	Constantly adjust car volume	64%	36%
7	Track Change requires adjustment	69%	31%

Question 4 & 5: asked the participants whether music was a distraction and whether radio interjects are too loud. The results show that 87% of the subject group are not distracted by music whilst driving. However, of the 13% remaining, a few participants indicated that music can be distracting during parking or with navigation.
The opinion that radio adverts/interjects are too loud reveals that 49% agree that adverts and interjects are too loud. This contrasts with the researcher's expectations by a small margin as 51% of the subjects in this study are not bothered by the loudness of adverts/interjects.

Question 6 & 7: asked the participants whether they constantly adjust the volume control on either radio or whilst listening to musical tracks on another sound system. From the participants' responses it can be seen that 65% feel the need to constantly adjust the volume when listening to the radio and 69% feel the change between musical tracks require volume readjustments. This meets the expectations of the researcher, supporting the concept that musical loudness both in radio and musical tracks require constant monitoring from the listener.

The next section highlights the loudness characteristics of the musical tracks used in the experiment.

4.2 Track Analysis

The loudness characteristics of each musical track used within the experiment as well as a discussion of the statistical P-value and LS Means between the two tracks are detailed below.

The following figures display an analysis of the musical tracks.

Figure 4.1: Pink Floyd Analysis

Figure 4.2: Saint-Saens Analysis

The TPL (fig. 4.1) is shown at -4.5 dBTP, meaning the full amount of headroom is not utilised by the Pink Floyd track. Conversely, Saint-Saens shows the full use of TPL set to -1.5 dBTP (fig. 4.2).

If a theoretical headroom value of 3 dB is added to the Pink Floyd track, the participants still would not have been exposed to a detrimental loudness level. Further audio processing was done to the Pink Floyd track post experimental stage to create an 'enhanced' track, proving that with an increase of 3 dB to -1.5 dBTP, the participants would not have been at any risk of hearing damage.

The enhancements were done using the Trans-X software (WavesAudio, 2016) to manipulate the transients of the Pink Floyd track, increasing the TPL from -4.5 dBTP to -1.6 dBTP (fig. 4.3). The newly enhanced Pink Floyd track was then played at the average level selected by the participants (25) and the highest value (32), as shown in Table 4.3.

Figure 4.3: Enhanced Pink Floyd Track

Using the DSP Mobile Analyzer (DSP-Mobile, 2012), the L_{eq}, with a slow response time was developed for both the average and highest values (fig. 4.4) as well as a peak dBA reading. The results comparing the original Pink Floyd against the enhanced version can be viewed in Table 4.3.

Table 4.3: Pink Floyd Peak dBA Comparisons

Input Value	Original PF (dBA)	Enhanced PF (dBA)
Average (25)	72.3	73.9
Highest (32)	85.6	86.1

Table 4.3 shows a minor difference of 1.6 dBA at the average listening level between the original PF and the enhanced PF track. The difference is even less at the highest playback level, showing a discrepancy of 0.5 dBA. At the average playback level for original PF, 72.3 dBA is more than 10 dBA higher than the background reference noise, which means the music loudness perception is perceived twice as loud as the background noise. Similarly, the result would be the same for the enhanced PF. At the highest playback level, both the original and enhanced PF tracks display values higher than 85 dBA. These peak values are however momentary rather than sustained and therefore would not pose any threat to the listener's hearing. The L_{eq} for the enhanced Pink Floyd track (fig. 4.4) is displayed below.

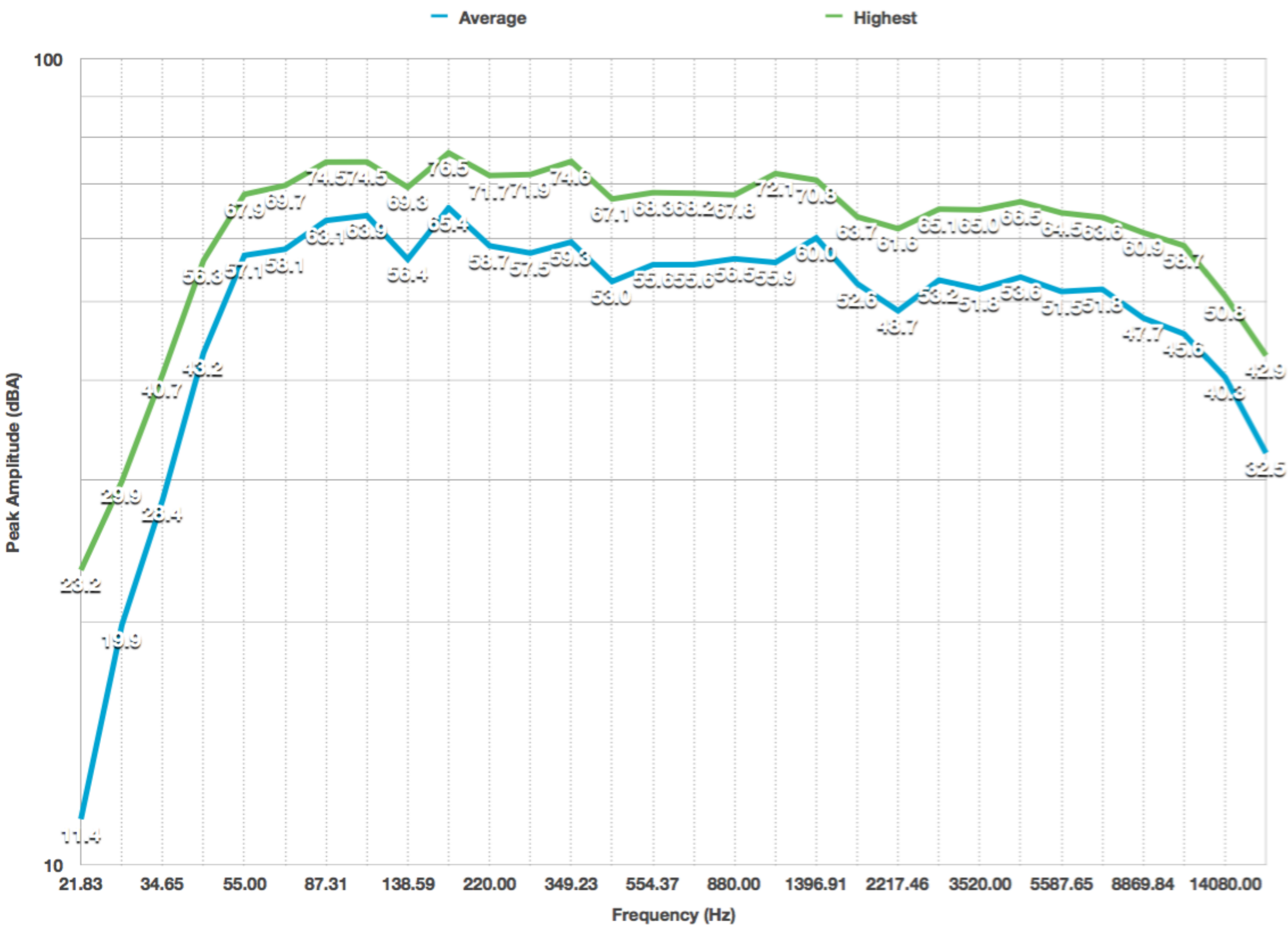


Figure 4.4: Enhanced Pink Floyd Frequency Plot

The variation is fairly constant between the frequencies 55 Hz and 11kHz in the overall loudness display across the 30 second segment (fig. 4.4). The constancy of the overall loudness portrays that even if the track utilises a fuller TPL, the enhanced PF does not generate a sustained harmful loudness level.

4.3 Musical Track LS Means

In addition to the loudness characteristics of each track, a LS Means graph was generated with the age group data gathered from the questionnaire, as well as with the peak decibel readings for each participant's car volume level. The data can be viewed in Table 4.4. The number of participants detailed in Table 4.4 differ from the previous data above as three participants chose not to state their age. Therefore the total number of subjects in this approximation equals 72 instead of 78. Furthermore, because of this slight discrepancy, the Peak dBA Mean value does not equal the average Peak dBA values for each track shown in Table 4.6. The mean values differ by a 1 dB in each case.

Table 4.4: Descriptive Statistics for Dependent Variables

Effect	Level of Factor	N	Peak dBA Mean	Peak dBA Std.Dev.
Total		72	70.91667	6.37511
Musical Track	Pink Floyd	36	73.33333	5.87070
Musical Track	Saint-Saens	36	68.50000	5.99714

The development of the LS Means graph is based on a null hypothesis, which in this case refers to both tracks being treated as equal. The Musical Track LS Means (fig. 4.5) displays a p-value measurement of 0.00001 which indicates strong evidence rejecting the null hypothesis, proving that in fact both tracks are dissimilar. The closer the p-value is to 1, the more likely it is that the null hypothesis can be proved correct. The p-value contributes perspective as to how reliable the experiment is. In order to aid in the monitoring loudness normalisation, the peak dBA mean shows how loud the participants enjoy listening to music. The standard deviation portrays a loudness range around the peak dBA mean whereby the loudness level is still accepted by the participants. The relationship between the two musical tracks (fig. 4.5) is shown below.

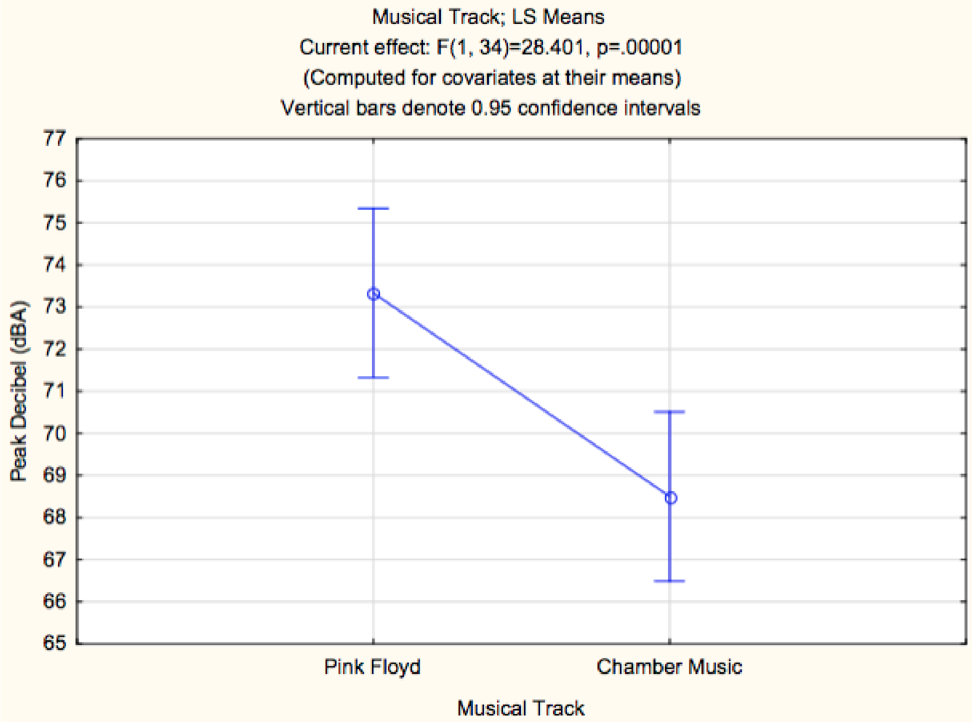


Figure 4.5: Musical Track, LS Means

The statistical LS means (fig. 4.5) shows that there is no overlap in the error bars, even with a large standard deviation. Each of the tracks present a clear and distinct difference, coupled with a p-value much lower than 5%, therefore there is insufficient evidence supporting the null hypothesis.

4.4 Participants' Loudness Selection

This section will detail the results of the participants' loudness selections. Table 4.5 represents the volume control value within the passenger vehicle adjusted by each participant during the experiment. The Renault Clio 2006 model shows a linear volume control scale that can be adjusted with buttons on the dashboard. Each participant's input was recorded as the number shown on the car's dashboard and consolidated to display the lowest, average and highest input values for each musical track.

Table 4.5: Participant's Car Value Input

Participant Inputs	Pink Floyd	Saint-Saens
Lowest	20	17
Average	25	23
Highest	32	32

The musical tracks were played at each value chosen by the participants, whilst using the DSP Mobile Analyzer to generate an L_{eq} at 1/3 octave bands as well as the peak dBA value. The measured peak dBA values for both tracks can be seen in Table 4.6.

Table 4.6: Car Values as Peak dBA

Participant Inputs	Pink Floyd	Saint-Saens
Lowest	63.1	56.7
Average	72.3	67.3
Highest	85.6	84.0

These peak dBA values give a more comparable representation of the loudness each participant preferred. The highest car value input for both tracks was set to 32, yet when this is converted into dBA, there is a more noticeable difference in loudness. The Pink Floyd track is louder at the same car value by 1.6 dBA. This proves the difficulty of designing a loudness algorithm to accurately adjust the loudness for all musical tracks, as the dynamic range of both tracks shown in this experiment cause a discrepancy in loudness playback. The data from Tables 4.6 and 4.7 is visually presented in Figure 4.6 and 4.7 respectively.

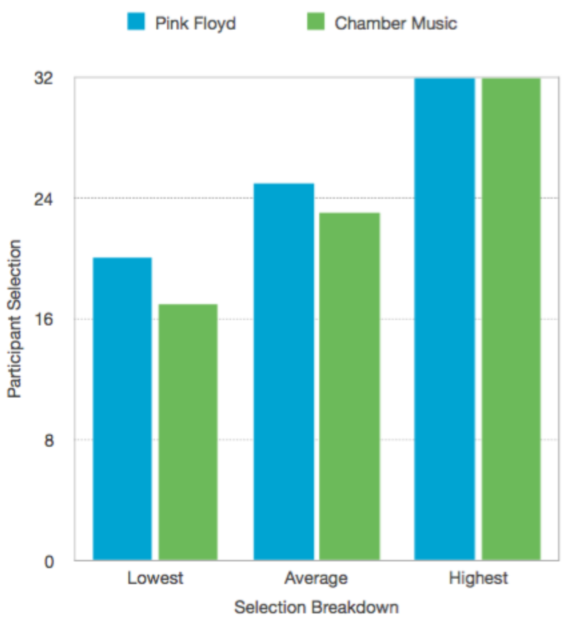


Figure 4.6: Car Values Selected by Participants

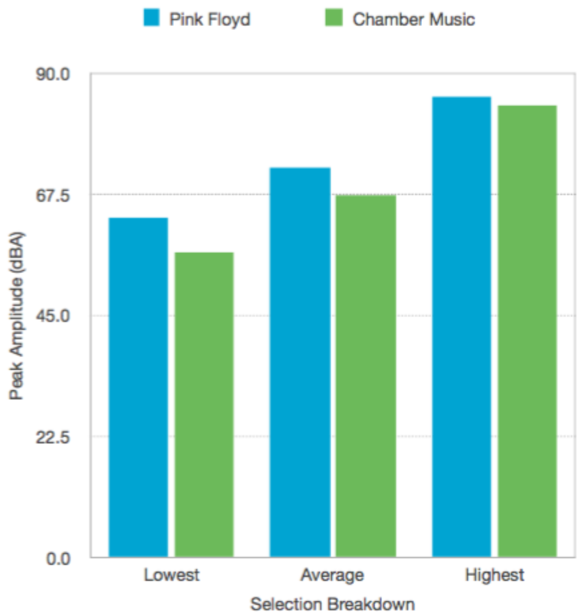


Figure 4.7: Decibel Values of Participant's Selection

The car value inputs for the lowest participant selection between the two tracks differ by three, with the average selection differing by two. Yet in a graphical comparison (fig. 4.6 & 4.7), the loudness discrepancy between the tracks is minimised. In order to understand what exactly each participant was exposed to, the peak values need to be accompanied by a frequency spectrum to show which frequencies are most prominent.

The frequency plot for both musical tracks (fig. 4.8 & 4.9), display the lowest, average and highest readings. The overall frequency exposure of the Pink Floyd track (as seen in fig. 4.8) remains constant. The peak amplitude raises from a minimum 48 dBA at 55 Hz and back down to 41.9 dBA around 11 kHz. This shows a constant amount of audible music across the spectrum. Conversely, the Saint-Saens (fig. 4.9) shows the lowest amplitude of 20.6 dBA at 55 Hz and 25.7 dBA around 11 kHz.

The replay experiment was carried out in the vehicle without the presence of the 60 dBA background noise. The L_{eq} profile for both the Pink Floyd and Saint-Saens tracks from which Figures 4.8 & 4.9 were graphed, can be seen in Appendix F and G respectively.

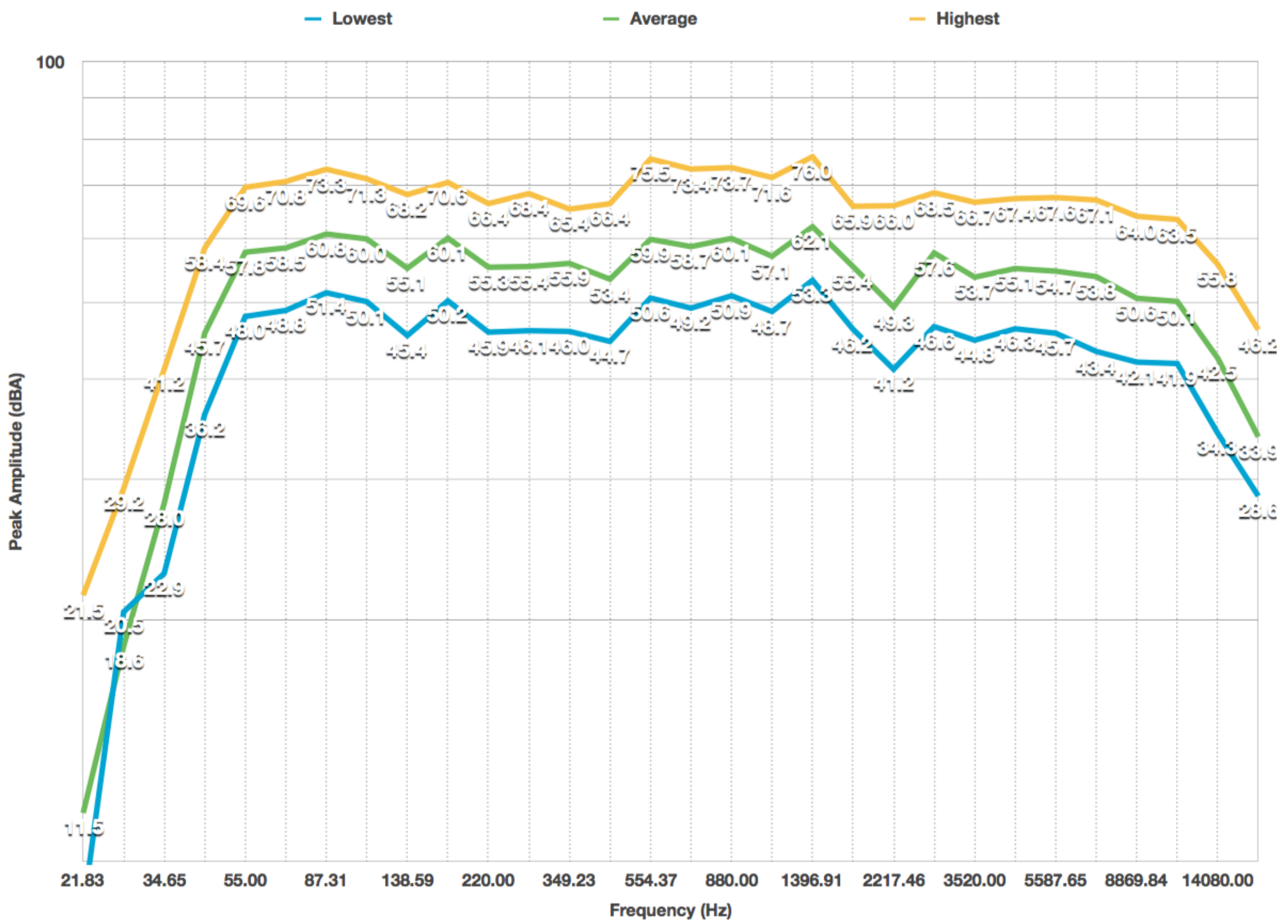


Figure 4.8: Pink Floyd Frequency Plot

The loudest frequencies present in the Pink Floyd track (fig. 4.8) are between 550 Hz and 1.4 kHz, where by the dBA reading peaks at 76 dBA. Similarly, in the Saint-Saens (fig. 4.9) within the same frequency range, the loudness levels peak at 75.8 dBA, displaying a minimal difference in comparison to the Pink Floyd.

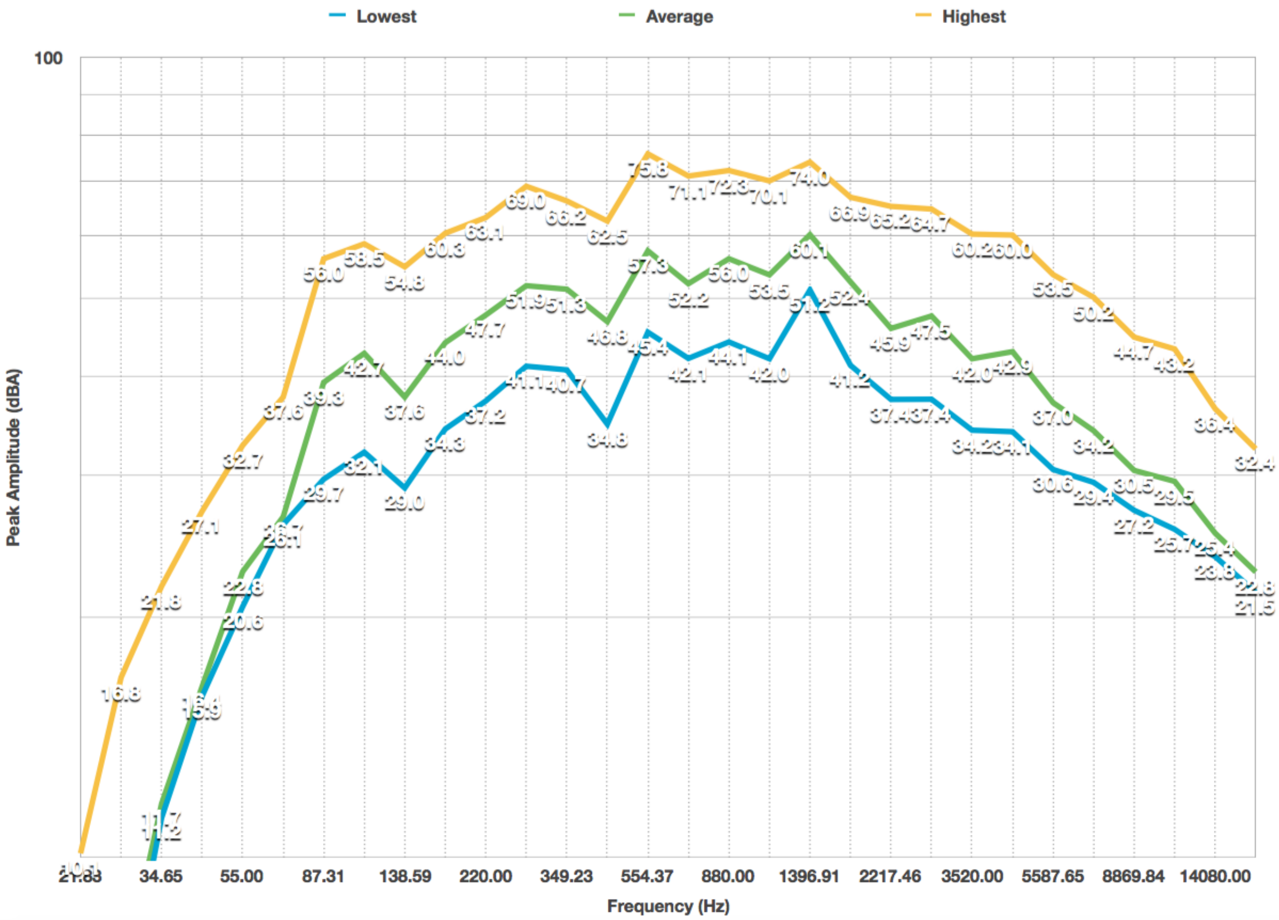


Figure 4.9: Saint-Saens Frequency Plot

The peak amplitude does not exceed 76 dBA (fig. 4.8 & 4.9) which is significant, as from an experimental point of view, none of the participants listened to either track loud enough to warrant hearing protection. From Chapter 2.1.2, it was noted that the European Legislation set the work environment noise level to 80 dBA (Howard & Angus, 2009:104). Therefore, even with the theoretical 3 dB increase to maximise the headroom of the Pink Floyd track as discussed in Chapter 4.3, the peak amplitude of the Pink Floyd would still be less than 80 dBA. The average peak amplitude of the Pink Floyd (fig. 4.8) shows a relatively consistent loudness output in the range of 50 - 60 dBA. The average peak amplitude of the Saint-Saens (fig. 4.9) shows a more varied loudness output, ranging from 22-60 dBA.

Due to the vast difference in average peak amplitudes across both musical tracks, the possibility of developing a loudness normalisation algorithm to precisely adjust the loudness cross-over between these two tracks would present a difficult amplitude compromise.

Chapter 5

Discussion

The results from this experiment are integrated with similar research and the potential implications of the results for the broadcast industry are discussed in this chapter.

5.1 Questionnaire Observations

The questionnaire, displayed in Appendix A, provided interesting demographic information regarding the participants and their background knowledge to loudness preferences. Question 1 pertaining to the age of the group of participants shows a wide age range from 19 to 38 years of age. In order to develop the most accurate representation of musical loudness preferences, an extended age range would be more ideal as the older individuals may present a completely different set of loudness results. As a loudness normalisation algorithm aims to monitor and adjust the loudness output for all radio transmissions, in order to develop a significant reference comfortable listening level, a fuller age range of participants would generate an age representative result.

Questions 6&7 asked whether the participants constantly adjusted the volume settings both on the radio and between musical tracks. This was necessary to ask as it shows directly whether individuals agree that listening to music over the radio requires adjustments. From the subject selection, 65% feel the need to constantly adjust the volume over the radio, whilst 69% feel the same toward musical tracks. Knowing that the majority of the test subjects are aware of the loudness spikes shows the need for the integration of loudness normalisation algorithms. Riedmiller & Robinson (2003:3) point out that the comfort zone is the loudness range of listener satisfaction. It is deduced from this subject group that the comfort zones of the listeners pertaining to musical playback both on and off radio is often disrupted by loudness inconsistencies. It is impossible to satisfy all the listeners, however utilising an average value recorded during this experiment could have provided a better overall basis to begin the loudness playback.

Question 8 showed the participants that there are readily available loudness normalisation methods in the public domain. It is interesting to see that 10% have utilised Apple SoundCheck and only 13% was aware of the Loudness Button in vehicles. The remaining 77% shared no awareness of the displayed loudness normalisation methods. This result did not come as a surprise, as the South African radio broadcasting industry does not currently implement, or show any signs of implementing loudness normalisation (Loots, 2016:19).

The lack of recognition toward the Loudness Button is understandable as some modern vehicles may not have a readily available button to initiate the normalisation, but rather hidden away in the settings (Jeep Renegade). The lacking awareness of the Apple SoundCheck is surprising due to the amount of iPhones used and easy accessibility through iTunes. The Apple SoundCheck function has been available in the iTunes settings since it's introduction in 2002 (Robjohns, 2014:1).

Lastly, Question 9 referred to the participants' knowledge of the audio format in which their music is stored and played. The results recorded were as expected, showing that 79% of the participant group utilised MP3s as their audio format. The collective responses for lossless audio format totalled 38%, which is understandable as several participants are practicing musicians/sound technicians, therefore understanding the higher fidelity audio formats. Skovenborg & Nielsen (2004:1) highlights that the audio format working in conjunction with the playback volume affects the subjects perception of loudness.

From the results of this participant group it could be established that the majority (79%) would give preference to lower fidelity audio purely due ease of access and familiarity.

5.2 Musical Loudness Integration

As stated in the previous chapter, the musical tracks were free from any destructive audio processing. In order to give the experiment an accurate resemblance to the ideal radio broadcasting simulation, the tracks needed to be high fidelity and processed in accordance to one of the loudness normalisation algorithms. The algorithm setting of choice was the EBU R128, with the loudness output set to -23 LUFS. The Saint-Saens had been recorded by Sun Studios at Stellenbosch University with the loudness output set to -23 LUFS as this is the studio formality. The Pink Floyd track was in FLAC format, which was converted to a wave file for the loudness output processing. Thereafter, both tracks in the form two wave files, 30 seconds in length were used for the experiment.

The choice of audio material is similar to Maempel & Gawlik (2009:2) as their experiment utilised six audio files, four popular, one classical and one spoken word. Their results included an average loudness level of 58 dBA over all their audio material. Furthermore, their experiment aimed at comparing five different processes as well as the unprocessed file, across six different genres Maempel & Gawlik (2009:5,8). The experimental procedure explained in Chapter 3, shows a similar use of audio content, utilising both popular and classical music to determine the desired loudness level, but the aims and the objectives of the studies were different so that direct comparisons are not possible. These studies of this nature highlight the need for further research.

During the experimental development, the methodology was initially designed to follow a similar approach as Riedmiller & Robinson (2003:3), whereby requiring participants to adjust a test track to a volume both higher and lower than a reference tone. The experimental procedure in this study aimed to develop a singular loudness value that could determine the optimal comfort value for radio listening rather than a full range of an individual's comfort zone. Furthermore, the difference in measurement approach allowed a loudness value for both the Pink Floyd and Saint-Saens to be acquired, whereas the results presented by Riedmiller & Robinson (2003:7,8) opted for an objective comparison between PPM, L_{eq} and VU meters.

In a study by Boley & Danner (2010:2,3), two musical tracks were presented to each participant, as was the case in the present study. Their experiment included several musical pairs in order to establish a comprehensive result. For the purposes of the current research, each participant was asked to adjust the musical tracks to their desired listening level, whereas Boley & Danner (2010:4) asked each participant to approximate the dynamic range of both musical tracks presented to them. Furthermore, Boley & Danner (2010:3,5) point out the evaluation of BS.1770 based algorithms to provide the best dynamic range option. The results of their research found insufficient data to support the use of an algorithm to estimate the perceived dynamic range. Whilst this may not directly influence the results obtained in the present study, testing the BS.1770 algorithms describes a similar approach to finding the solution of loudness algorithm integration to determine the dynamic range and preferred loudness output level.

This research aimed to extract the participants' immediate selection for a preferred loudness listening level. Therefore in addition to the varied choice material, the musical tracks were kept to 30 second segments. Repeating the experiment with more musical pairs in this case would give the participant more time to adjust their preference with each passing musical track. This would arguably defeat the purpose of determining their initial loudness preference.

The initial results shown in Table 4.5 display the volume selection on the car by each participant, allowing the researcher to distinguish the linear difference in level preference for each musical track. However, for a more comprehensive analysis Table 4.6 is used to give a dBA reading for each track. The results show that the average loudness for the Pink Floyd track is 72.3 dBA, which is 12.3 dBA higher than the interior noise reference level (60 dBA). This means that from a psychoacoustic loudness point of view, the average sound level is played double the perceptible loudness level of the interior reference level. The highest dBA value for each track shows to be around 25 dBA higher than the reference tone, meaning the perceptible loudness level is about five times louder. Whilst this arguably presents a dangerous level, these peak dBA values are not sustained throughout the audio track but rather instantaneous loudness readings.

Howard & Angus (2009:104) point out the change in European Legislation for the regulation of loudness levels in the working environments through the implementation of a First Action Level. This value is set to 80 dBA, whereby any individual exposed to excess noise levels greater than 80 dBA should consider wearing hearing protection or attenuating the volume. It can be seen from Table 4.6 that both the Pink Floyd and Saint-Saens produce dBA readings greater than 80 dBA, but not peaking much above 85 dBA. The average listening levels were between 67 dBA and 73 dBA, vastly softer than the First Action Level. The few values peaking above 85 dBA are not sustained as proved by both L_{eq} graphs for each track (fig. 4.8 & 4.9) respectively.

In the original Pink Floyd L_{eq} frequency plot (fig. 4.8), the green line represents the average dBA output showing a slight fluctuation either side of 55 dBA. This fluctuation means that the average individual was exposed to between 49.3 - 62.1 dBA over the entire Pink Floyd track. The peak decibel value for the average participant from Table 4.6 shows the exposure to 72.3 dBA, significantly higher than the sustained loudness output. Furthermore, the yellow line (fig. 4.8), represents the highest input value by some participants, showing a fluctuation between 66 - 76 dBA over the 30 second segment, 9.6 dBA lower than the peak value in Table 4.6.

The green line in the Saint-Saens L_{eq} frequency plot (fig. 4.9) displays the average input with loudness readings fluctuating markedly in comparison to the Pink Floyd readings (fig. 4.8). However, the majority of the track lies between 39.3 - 60.1 dBA, again much lower than the peak value of 67.3 dBA from Table 4.4. The yellow line (fig. 4.9), displays the highest input by some of the participants presenting a fluctuation between 56.0 - 75.8 dBA, again, 8.2 dBA lower than the respective peak value in Table 4.6.

5.3 Enhanced Pink Floyd

The data presented in Chapter 4.3 indicated that the Pink Floyd track utilised up to -4.5 dBTP, whereas the Saint-Saens went up higher to -1.5 dBTP. This shows that the Pink Floyd track has a more compressed dynamic range than the Saint-Saens track. If the Pink Floyd track had the full use of its dynamic range, the listener's hearing would still be safe to listen at their preferred levels. In the present research, the Pink Floyd track was manipulated (fig. 4.3) to bring the TPL to -1.5 dBTP, but this enhanced version of the track was not used in the experiment. The aim of the enhanced version was to highlight the safety of the theoretical 3 dB headroom increase.

In Table 4.1, the comparison between the Original and Enhanced Pink Floyd tracks shows only a small increase of 1.6 dBA for the average and 0.5 dBA at the highest level. Yet the frequency plot of the enhanced version (fig. 4.4) shows that the L_{eq} does not exceed 76.5 dBA, well within the safety of the European Loudness Legislation. Furthermore proving that even with a theoretical headroom increase of 3 dB to the Pink Floyd track, the listeners' hearing is still within a safe loudness level.

The relationship between the original and enhanced Pink Floyd tracks (fig. 5.1) are presented below. The purple and red line represent the original Pink Floyd track detailing the Peak Amplitude in dBA, whilst the light blue and orange show the enhanced Pink Floyd track utilising the -1.5 dBTP level. Unexpectedly, the enhanced Pink Floyd shows both dips and raises above the original, rather than a pure increase in Peak Amplitude across the board. This result is crucial as it shows that even with the headroom increase, participants would be at no more risk with the enhancement than the original if played at the same level. Consequently, demonstrating that over radio broadcast, popular music played with either a full or compressed dynamic range will produce a similar output loudness.

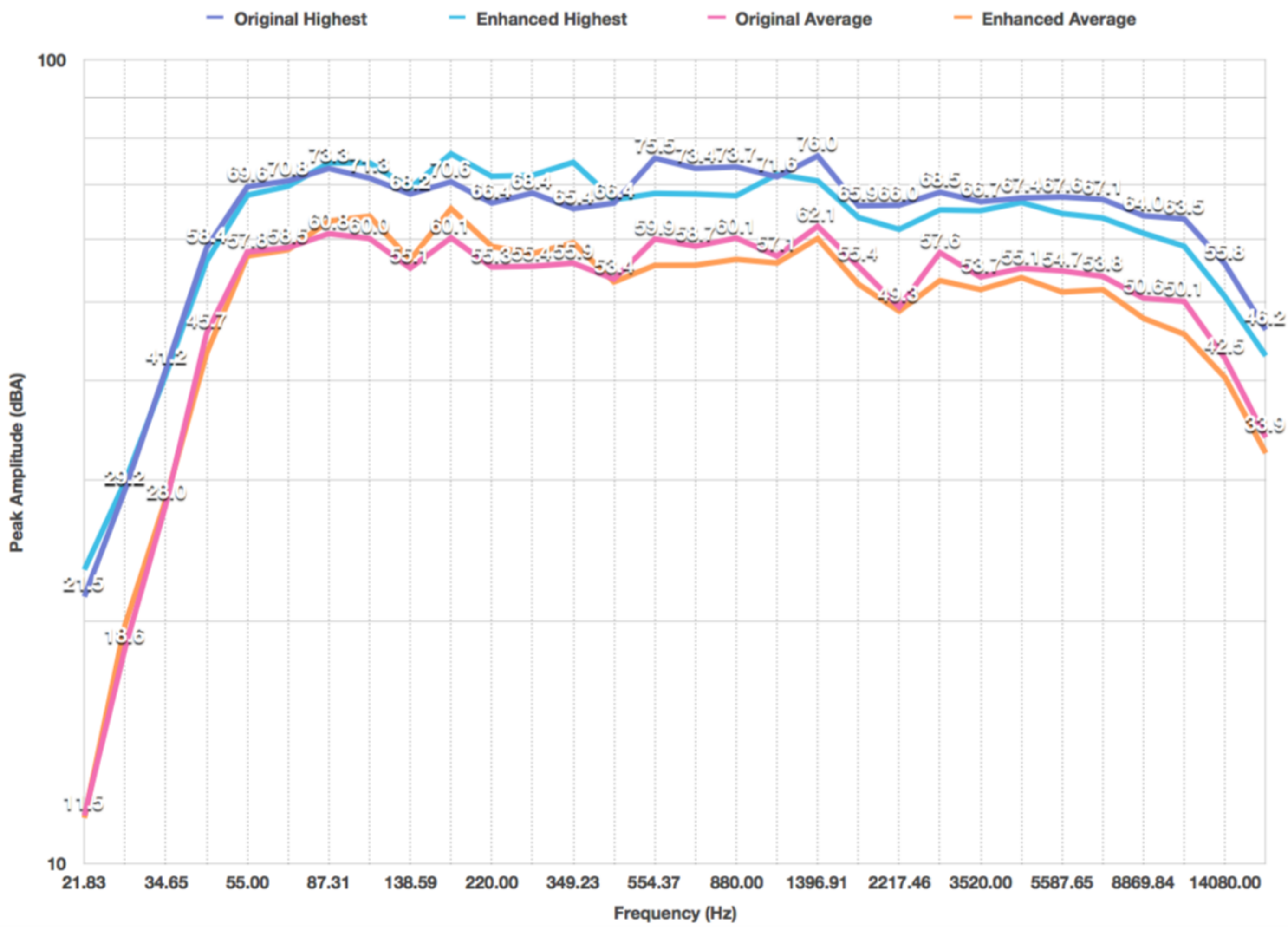


Figure 5.1: Original vs Enhanced Pink Floyd

From the point of view of radio broadcasting, Hadi (2010:10) states that dynamic range music broadcast depends on the listening environment's noise floor. Whilst the noise floor in this experiment was set to be 60 dBA, the average listening level for both Pink Floyd and Saint-Saens was measured at 72.3 dBA and 67.3 dBA respectively. As a consequence, it shows that even with dynamic tracks, an adequate loudness level deemed comfortable by the participants is achievable over the loudness of the vehicle's noise floor.

Furthermore, through the testing of the enhanced Pink Floyd track against the original, the dBA values that a listener would be exposed to shows a minimal difference in loudness (fig. 5.1). The implications here reveal that in the case of this experiment, the switch from a popular track with a compressed dynamic range and minimal transients, to the same track with a fuller dynamic range and more elevated transients will result in a slight loudness output reduction as represented by the blue and purple lines.

The next section will discuss the statistical LS Means of the Pink Floyd agains the Saint-Saens.

5.4 Statistical LS Means

The statistical LS Means provides insight as to whether the same result could be obtained should the experiment be repeated again with the same audio stimuli and subject size. With the p-value at 0.00001, there is a high probability that the experiment will produce the same result. In other words, rejecting the null hypothesis. Lund (2006:59) points out the subjectivity of listeners depends on frequency contents, duration and SPL. This experiment managed to control the duration of the audio tracks and processing, ensuring the same frequency output from the vehicle's radio for each listener. Therefore, the SPL and frequency content experienced by each listener would fluctuate based on the anatomical structure of their pinnae.

The peak dBA mean provides insight into the comfort level chosen by the average number of participants, whilst the dBA standard deviation shows the relationship between the participants compared to the average selection. The loudness range from the volume selection of the participants (fig. 4.5) has a small variation of about 2 dBA either side of the mean dBA over both tracks. The significance of these error bars portray the loudness range for the participants, whilst displaying the connection with the preferred loudness level of both musical tracks.

Furthermore, the experiment ensured consistency across the audio format output as both the Pink Floyd and the Saint-Saens were played through Audacity in a Wave Audio file type. According to Skovenborg & Nielsen (2004:1) the loudness perception by each individual depends both on the volume, but also the audio format of the track. Since the experiment aimed to acquire the participant's loudness preference, the variables pertaining to the frequency content and audio formatting for each track had to remain consistent for each subject's participation.

5.5 Result Impact

This section will incorporate the loudness results for each track and present suggestions to SA radio stations. As stated previously, Loots (2016:19) points out that DSTV is the only broadcasting service to incorporate loudness normalisation in accordance with the EBU R128. Neither the SABC nor local radio stations incorporate any loudness normalisation algorithms or BS.1770 based metric monitoring systems.

In Question 8 of the Questionnaire, participants were asked whether they recognise any of the loudness algorithms. With several of the participants in the subject pool working for a local radio station in Stellenbosch, none of them recognised ReplayGain. With the overall move away from broadcasting compressed audio to dynamic audio, the local radio stations could utilise independently developed software to aid in loudness monitoring. In a similar experiment carried out by Nygren (2009:2), the radio station under observation utilised audio software called Awave Audio, utilising ReplayGain for radio broadcast which was implemented into Swedish Radio.

Since South African radio broadcast is still far from incorporating loudness normalisation, the use of metering systems such as MasterCheck may help the production engineers produce the on-air musical tracks closer to the BS.1770 standard. NugenAudio (2014:4) highlights that the MasterCheck meter allows for the producers to maximise the additional headroom of a given track without the confusion of loudness mismatches. Furthermore, the use of MasterCheck will allow for the monitoring of groups of musical tracks, making musical transmission over radio at a standardised loudness level.

The loudness results discussed in Chapter 4, show that the average individual sets the musical comfort level to be at 25 and 23 on the car volume control for the Pink Floyd and Saint-Saens respectively. From the Peak dBA readings for each volume setting, the average Pink Floyd and Saint-Saens loudness output peaks at 72.3 dBA and 67.3 dBA. These values may indicate the most appropriate loudness levels for passenger car radio listening.

These appropriate loudness levels may be useful for a device controlling the compression levels at the users end. Should the radio transmit high fidelity and dynamic audio, the device could adjust the levels of compression to suit the listening environment. In the case of a passenger car, the device could integrate the peak loudness levels for the average listener’s comfort level and adjust the audio compression appropriately. This would ensure the music could be heard above the engine noise without any distracting adjustments from the driver.

The next section will draw conclusions for this study.

Chapter 6

Conclusion

The aims of this research study were firstly, to determine the preferred loudness playback level for passenger car drivers whilst listening to two audio tracks against background noise and secondly, to determine the optimal loudness range for radio transmission.

In terms of the first aim, the results obtained show that at the peak dBA reading, the average participant selected 72.3 dBA (Volume level 25) for the Pink Floyd, with a selection of 67.3 dBA (Volume level 23) for the Saint-Saens. At the average loudness level, the Pink Floyd selection fluctuation 5 dBA either side of 55 dBA and between 40 - 60 dBA in the Saint-Saens. The European Legislation set the requirement for hearing protection to be worn in excess of 80 dBA. Therefore, due to the reference level being set at 60 dBA, the subjects increased the loudness level of the music only marginally higher to be heard above the surrounding background noise.

The outcome loudness levels indicated by the participants in this study show that their selections did not exceed 80 dBA. Instead, the loudness levels chosen were within a safe listening level, therefore potentially not exposing participants to hearing damage within the passenger car environment.

In terms of the second aim, to determine the loudness range in the passenger car environment, it can be concluded from the results that the L_{eq} amplitude levels for both tracks did not exceed 76 dBA. However, the instantaneous loudness levels peak at 85.6 dBA and 84 dBA for the Pink Floyd and Saint-Saens respectively. These two readings are only from the minority of subjects listening to the music at the highest volume level. The loudness values for the average listening levels give an indication of an output level for radio broadcasting. With dynamic audio returning to radio broadcast, the preferred listening levels for each of the genres discussed in this study may provide insight into an acceptable loudness range for musical broadcast.

Testing compressed audio against dynamic range reliant audio shows that within the passenger car environment, the track with a smaller dynamic range (Pink Floyd), was in fact preferred at a higher loudness level than the Saint-Saens. In support of the hypothesis, both tracks were increased high enough from the average subject in order to be enjoyed at a comfortable listening level. The addition of a 3 dB theoretical headroom addition to the dBTP of the Pink Floyd proves that if listeners were exposed to the enhanced Pink Floyd track, they would still be within a safe listening loudness range.

This study details the importance of loudness normalisation algorithms, as well as how they may be incorporated into radio broadcast. The EBU R128 aims to be the provider of a worldwide loudness normalisation algorithm. With independent developers and software plugins based on the BS.1770, loudness normalisation is more easily accessible for both broadcasters and individual listeners. The development of loudness descriptors shows a progression towards better control and monitoring of musical loudness levels. In the near future, with the development of descriptors in conjunction with the normalisation algorithms, radio broadcasters will have a detailed method of monitoring and appropriately controlling the output loudness of musical tracks.

Within the context of the present research study, it can be pointed out that as loudness normalisation has not been implemented fully across radio broadcasting, the drivers may ignore the essential need to readjust the volume control resulting in listening to extended loud musical broadcasts. This prolonged exposure may lead to hearing damage in the long term.

With the return of dynamic audio into radio broadcast, this study may provide insight into a preferred loudness playback value for these dynamic musical recordings. In order to control musical loudness outputs within the music industry, the integration of loudness normalisation by radio broadcasters, or at least loudness monitoring, should therefore be a mandatory effort from both the public and broadcasting stations.

Chapter 7

Recommendations for Further Research

Based on the conclusions presented in the previous section of this thesis, this chapter will present recommendations for future research and also highlight the limitations of the present research.

7.1 Recommendations for Further Research

Based on the results and conclusions of the present study, the following suggestions are made for future research:

Listening Pleasure: Due to the dynamic range compression across the radio broadcasting industry, there is an overall loss in listening pleasure. Developing a qualitative assessment of the perceived pleasures of compressed and dynamic music over radio broadcast could reveal the extremity of damage caused to music through dynamic range compression. The subjects could be expanded to audiophiles for a more accurate description of dynamic reduction and listening pleasures.

Through the incorporation of musically inclined individuals, a better understanding of the loss of both the composer's dynamics and decorated notes from classical music can be highlighted. This could drive motivation for the implementation of loudness normalisation algorithms for music over radio broadcast.

Loudness Regulation: Through regulating the loudness output, the driver would no longer be required to adjust their volume control constantly. This would result in an increase of concentration on the road, less likely to create a hazardous situation for the driver. Presenting both a perfect or imperfect environment to a subject pool may bring the attention to incorporating simpler systems of loudness normalisation to ease the problem. Since the implementation of loudness normalisation in radio broadcast will take years, alerting the drivers to the environment of musical enjoyment free from distraction may help motivate the process.

This could be done through presenting an extended music mix containing various musical tracks and genres produced at -23 LUFS, asking the subjects to input the playback volume setting. The subjects could then note down their experience and whether there was any cause for distraction or readjustment, bringing their attention to the issue of inconsistent loudness playback. The psychological attributes that would accompany this study may be considered beyond the scope of the field of psychoacoustics.

Loudness Descriptors: With the development of loudness descriptors, a visual analysis of the objective characteristics of the musical tracks can be displayed. Utilising this prototype software would give a new perspective to the loudness levels drivers in a passenger vehicle are exposed to. Repeating the experiment with a variation in loudness algorithms, coupled with the visual display from the loudness descriptors may provide insightful information to which method of loudness normalisation is best for music listening over radio broadcast.

Hearing Loss: With the increased awareness of the impact of listening to loud music, further research into listening to loud music in the car environment may be valuable. Exposure to loud music over long periods of time may cause hearing damage, nevertheless individuals choose to listen to music at moderate to high volume levels. Due to the fluctuation in musical loudness, the volume levels may be increased in an attempt to hear the subtleties within dynamic musical tracks in comparison to the hypercompressed tracks. In the case of a passenger car, tying in the distraction of adjusting playback volume, drivers may opt for leaving the volume control on a louder setting, thus negating the distraction.

An experimental design may take the form of a dummy head positioned in the driver's seat of a car, used to monitor the effects of long term loud music exposure. Using the highest volume setting from this experiment and prolonging the exposure over a time period of several hours, may prove the damaging effects to hearing through the analysis from the dummy head.

7.2 Present Research Limitations

The experiment encountered several limitations that could be improved should the experiment be repeated.

Subject Age Group and Gender: The subject age group for this experiment covered a 19 year age range, with the youngest participant at 19 and the oldest at 38. The researcher approached fellow university students and staff members to participate in the research which meant the exclusion of older possible participants. The incorporation of individuals at least up to the age of 65 would show a wider set of loudness preferences, likely at both the louder and softer extremes depending on the age of participant.

In addition to expanding the age range, comparing an equal number of male and female participants may show whether gender plays a role in an individual's loudness preference.

Sample Size: The experimental procedure included 39 participants. Whilst this number provides an understanding of loudness preferences of listeners in a car environment, repeating the study with a large sample of the population may yield more specific insights. Such results may also provide necessary pointers to the broadcast industry in terms of an appropriate average loudness playback level.

Questionnaire Design: A questionnaire of nine questions was used to obtain demographic information regarding participants. However, retrospectively, several questions provided minimal value information to the study. The questions pertaining to the noise related habits of each participant show the amount of loud music exposure which was initially thought to mirror their loudness comfort level, however was later found to be irrelevant for the scope of this thesis.

The questionnaire should rather have included more Likert questions pertaining to how the listener perceived the musical tracks. For example: On a scale of 1 to 10, how well are the delicate notes of the Saint-Saens perceived over the background noise. This would provide insight into how well dynamic music is received in the imperfect listening environment of the passenger vehicle. A pilot study involving the questionnaire would have been valuable to ensure that all questions contributed exactly what is needed for the research.

Musical Material: It is felt that the musical material was limited to a pre-selected choice of two musical tracks. Whilst the use of only two tracks provided an immediate and instantaneous response for loudness preference from each participant, the incorporation of more musical pairs within the same genre would create a comparison between the tracks of both the popular and classical genres.

In addition, a greater variation of musical genres would aid the development of a preferred loudness level as radio broadcasting is not limited to popular rock and classical music. The inclusion of a spoken word track, as well as music from an electronic genre would give additional value to widely listened to genres.

Background Noise Simulation: The present experiment utilised pink noise to simulate sufficient background noise surrounding the vehicle. The pink noise was chosen as it simulates the sound of moving traffic and allows for the control over the amount of noise each subject would experience, keeping the noise variable constant. Upon reflection, recording an extended sound file of moving traffic, could have provided more specific background noise.

Similarly, recording the engine noise at a constant speed could otherwise provide better insight into the real world environment each driver would be subjected to whilst listening to the radio.

By attending to the limitations encountered in the present study, a simulation for listening to music over radio broadcast built using a real world environment would create a pratical approach, which may contribute to the precision of the research.

Appendices

Appendix A

Investigation Questionnaire

Driver's Loudness Preferences **Participant Number:____**

Thank you for volunteering to aid me with my investigation into the driver's preferences in loudness of musical recordings. What we are about to do shouldn't take longer than a couple minutes. The questions below will help me to build a participant demographic. If you have any questions of your own, please feel free to ask.

1. Please state your age: _____

2. Please tick each category that best describes your noise related habits

Domestic	Occasional music in your car		Often listen to music in your car		Always listen to music in your car	
Social	Occasionally go to clubs/bars		Often go to clubs or bars		Go to clubs or bars at least 2-3 times a week	
Hobby	Occasionally go to live concerts/ festivals		Often attend live concerts and festivals		Attend concerts and festivals multiple times a year	

3. How many hours per day are you exposed to moderate to loud music?

0-2hrs 3-4hrs 4-6hrs 6hrs+

4. Do you find music in your car to be a distraction?

Yes No

5. Do you feel that radio interjects/adverts are too loud?

Yes No

6. Do you have to constantly adjust the volume in your car when listening to the radio?

Yes No

7. Do you feel that switching between musical tracks often require volume readjustments?

Yes No

8. Have you utilised any loudness normalisation methods? (To keep the volume level across your music constant)

Apple SoundCheck Loudness Button in Cars ReplayGain None

9. What format do you normally use for your music?

FLAC Wave MP3 AAC AIFF Don't know

Appendix B

Question Motivation

Below is a brief outline of each question followed by the motivation for it's choosing.

1. **Please state your age.**
   This question provides the researcher with an age range of all participating subjects.

2. **Please tick each category that best describes your noise related habits.**
   This question gives the best description of how often each participant listens to music in their vehicle, attends bars and clubs as well as festivals. This is of importance as it gives the researcher an indication of loudness preference with regard to the participants' hobbies. Since the investigation focuses on loudness levels in a passenger vehicle, it is necessary to note how many individuals listen to music in their vehicle. This gives rise to the urgency of loudness algorithm implementation. Should every participant always listen to music in their vehicle, the use of loudness normalisation becomes a necessity for the comfort of each driver.

3. **How many hours a day are you exposed to moderate to loud music?**
   The question of how many hours each participant is exposed to loud music on a daily basis ties in with Q2 as it gives an indication of how the subject may enjoy their loudness preference. Moreover, should the subject be exposed to loud music over long periods of time, they may be prone to listening fatigue and as a result, listen to the music in their car at a higher volume.

4. **Do you find music in your car to be a distraction?**
   Asking the participants whether they find music to be a distraction provides insight as to whether their choice in musical loudness is associated with breaking their concentration whilst driving.

5. **Do you feel that radio interjects/adverts are too loud?**
   This question is designed to bring the participant's attention to whether or

not interjects and adverts on the radio are too loud. This provides insight as to whether the subject is aware of the loudness change between musical track and advert. Should the participant feel that the adverts are too loud, it shows that advert loudness can present a distraction to the driver.

6. **Do you have to constantly adjust the volume in your car when listening to the radio?**
   Question 6 ties in with Q5 as it shows whether the participant notices any loudness variation over the radio. Furthermore, should the participant answer: 'Yes', it shows that the participant acknowledges the annoyance and attends to it.

7. **Do you feel that switching between musical tracks often require volume readjustments?**
   Question 7 brings the attention of the participant to the difference in loudness levels between musical tracks in addition to the music-advert loudness discrepancy. Should participants answer with 'Yes', it alerts the researcher to the fact that they are more likely to attend to the matter of volume misalignment.

8. **Have you utilised any loudness normalisation methods?**
   This question is designed to show how many participants are aware of the loudness normalisation algorithms that can be accessed freely to attend to the music loudness discrepancies.

9. **What format do you normally use for your music?**
   The final question is designed to show how many of the participants utilise high fidelity audio when listening to music. This is important as the use of more dynamic and compressed audio may have an impact on the preference of loudness when played in their vehicles.

Appendix C

Experiment Description

Title of the Research Project: An Investigation into Passenger Car Drivers' Preferences in Loudness between Dynamic and Compressed Musical Recordings.

Ethics Reference Number: SU-HSD - 003089

Researcher: Mark Stobbart

Address: Music Dept, Neethling Street, Stellenbosch University, 7602

Contact Number: 0766129866 (Investigator)

Dear Fellow Student,

My name is Mark Stobbart and I am investigating music loudness levels within a passenger car. I would like to invite you to participate in a research project titled, "An Investigation into Passenger Car Drivers' Preferences in Loudness between Dynamic and Compressed Music Recordings".

Please take some time to read the information presented here, which will explain the details of this project and contact me if you require further explanation or clarification of any aspect of the study. Also, your participation is **entirely voluntary** and you are free to decline to participate. If you say no, this will not affect you negatively in any way whatsoever. You are also free to withdraw from the study at any point, even if you do agree to take part.

This study has been approved by the **Humanities Research Ethics Committee (HREC) at Stellenbosch University** and will be conducted according to accepted and applicable national and international ethical guidelines and principles.

The purpose of this investigation is to measure the desired loudness level for music playback within a passenger car. The collection of data will be as follow:

- Seat you in the driver seat of a passenger car.
- Play you a couple musical tracks to evaluate your loudness preference through adjusting the volume control.

- Fill out a short questionnaire pertaining to: Your age and Noise related hobbies
- Participants may choose not to answer certain questions and still remain in the study.
- The potential for negative effects or risks are minimal.
- Benefits of participation: Allows the investigator to develop a value for the preferred loudness level when listening to music in a passenger vehicle. The value is important for the normalisation of audio across broadcasting stations and musical tracks, as it portrays the playback loudness level deemed comfortable by drivers.
- To ensure your anonymity, the participant will be assigned a number and the raw data will be discarded upon completion of the thesis.
- The investigation is entirely voluntary, however should you feel uncomfortable or wish to discontinue your participation, you may do so without any negative consequences.

If you have any questions or concerns about the research, please feel free to contact either the investigator or supervisor at:

- Mark Stobbart (Principal Investigator)
- E: mstobbart@gmail.com
- T: 0766 129 866
- Gerhard Roux (Supervisor)
- E: groux@sun.ac.za
- T: 021 808 2138

Rights of Research Participants: You may withdraw your consent at any time and discontinue participation without penalty. You are not waiving any legal claims, rights or remedies because of your participation in this research study. If you have questions regarding your rights as a research subject, contact Ms Malene Fouche (mfouche@sun.ac.za; 021 808 4622) at the Division for Research Development. You have right to receive a copy of the Information and Consent form. **If you are willing to participate in this study please sign the attached Declaration of Consent and Return it to the Investigator**

Yours sincerely
Mark Stobbart (Principal Investigator)

Appendix D

Participant Declaration of Consent

Declaration by Participant

By signing below, I ...., agree to take part in a research study titled An Investigation into Passenger Car Drivers' Preferences in Loudness between Dynamic and Compressed Musical Recordings and conducted by Mark Stobbart.

I declare that:

- I have read the attached information leaflet and it is written in a language with which I am fluent and comfortable.
- I have had a chance to ask questions and all my questions have been adequately answered.
- I understand that taking part in this study is **voluntary** and I have not been pressurised to take part.
- I may choose to leave the study at any time and will not be penalised or prejudiced in any way.
- I may be asked to leave the study before it has finished, if the researcher feels it is in my best interest, or if I do not follow the study plan, as agreed to.
- All issues related to privacy and the confidentiality and use of the information I provide have been explained to my satisfaction.

Signed at (place) .... on (date) .... 2016.

.... **Signature of participant**

Signature of Investigator

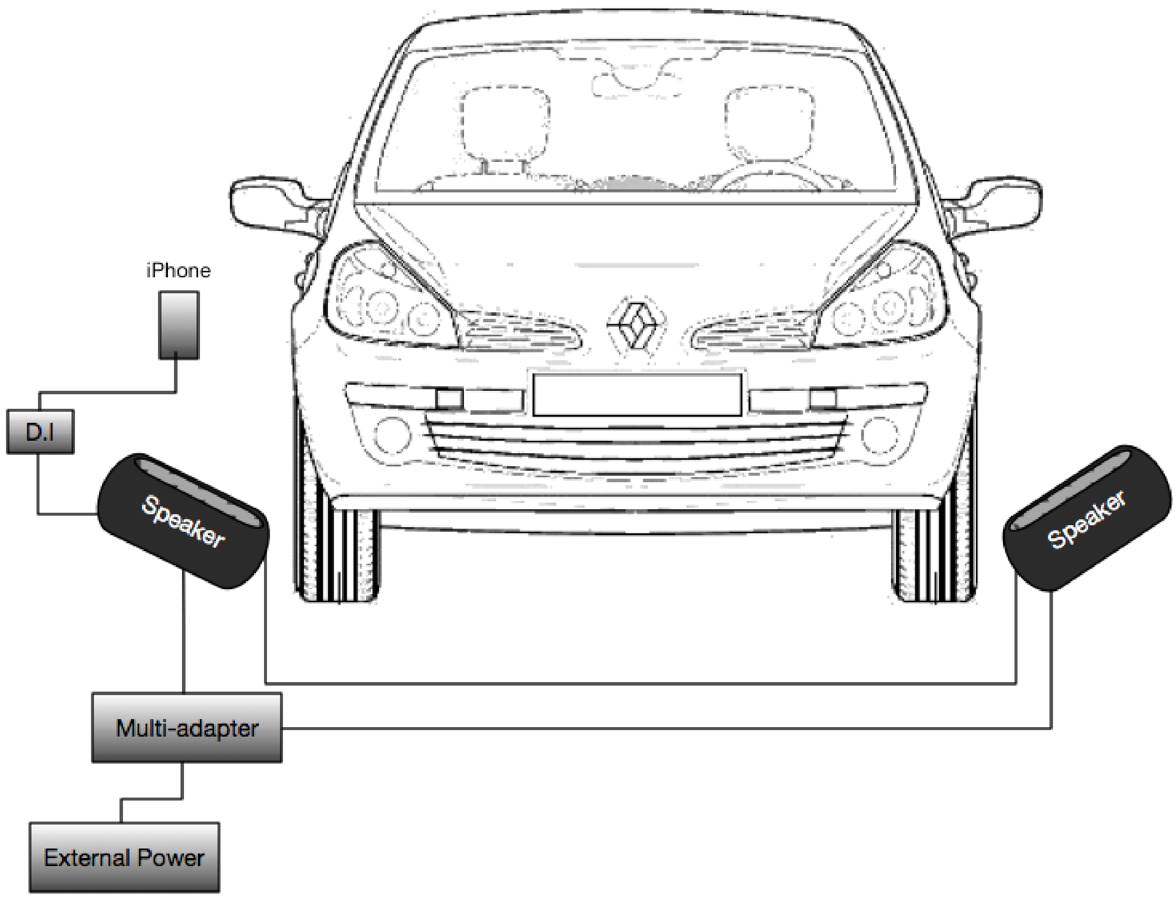
I declare that I explained the information given in this document to ...... (name of the participant). (He/she) was encouraged and given ample time to ask me any questions. This conversation was conducted in (Afrikaans/English/Xhosa/Other) and (no translator was used/this conversation was translated into ...... by ......).

.... **Signature of Investigator**

.... **Date**

Appendix E

Experimental Setup



Renault Clio diagram provided by CarBlueprints (2008).

Appendix F

Pink Floyd Frequency Plot Data

Pink Floyd: Frequency - Amplitude, CSV Data			
Frequency (Hz)	Amplitude Values (dBA)		
	Lowest	Average	Highest
21.83	8.69	11.49	21.51
27.50	20.50	18.58	29.25
34.65	22.90	28.02	41.23
43.65	36.21	45.73	58.44
55.00	47.99	57.75	69.61
69.30	48.83	58.46	70.79
87.31	51.38	60.85	73.34
110.00	50.08	60.00	71.33
138.59	45.42	55.14	68.16
174.61	50.20	60.13	70.60
220.00	45.87	55.31	66.42
277.18	46.06	55.41	68.37
349.23	45.97	55.92	65.37
440.00	44.66	53.43	66.40
554.37	50.61	59.93	75.53
698.46	49.15	58.69	73.37
880.00	50.94	60.08	73.68
1108.73	48.68	57.08	71.61
1396.91	53.26	62.14	76.02
1760.00	46.25	55.40	65.91
2217.46	41.23	49.31	66.02
2793.83	46.61	57.63	68.51
3520.00	44.82	53.74	66.68
4434.92	46.32	55.09	67.40
5587.65	45.73	54.68	67.59
7040.00	43.40	53.83	67.11
8869.84	42.08	50.57	64.04
11175.30	41.91	50.11	63.46
14080.00	34.28	42.55	55.77
17739.69	28.62	33.94	46.20

Appendix G

Saint-Saens Frequency Plot Data

Saint-Saens: Frequency to Amplitude, CSV Data

Frequency (Hz)	Amplitude Values (dBA)		
	Lowest	Average	Highest
21.83	-1.19	0.43	10.12
27.50	6.12	6.83	16.79
34.65	11.21	11.68	21.84
43.65	15.91	16.35	27.10
55.00	20.59	22.77	32.75
69.30	26.07	26.66	37.63
87.31	29.71	39.26	56.05
110.00	32.12	42.69	58.51
138.59	28.98	37.64	54.75
174.61	34.33	44.01	60.31
220.00	37.22	47.66	63.10
277.18	41.13	51.86	69.04
349.23	40.70	51.34	66.19
440.00	34.81	46.76	62.47
554.37	45.36	57.33	75.79
698.46	42.07	52.16	71.11
880.00	44.10	56.04	72.25
1108.73	42.03	53.51	70.12
1396.91	51.21	60.07	74.01
1760.00	41.24	52.44	66.89
2217.46	37.38	45.86	65.17
2793.83	37.40	47.53	64.66
3520.00	34.21	42.03	60.16
4434.92	34.07	42.89	60.00
5587.65	30.55	37.03	53.54
7040.00	29.44	34.17	50.16
8869.84	27.17	30.47	44.73
11175.30	25.72	29.53	43.23
14080.00	23.77	25.43	36.39
17739.69	21.46	22.75	32.43

List of References

AcousticFields. 2016. White Noise Definition Vs. Pink noise. Accessed: October 2016. Available at: http://www.acousticfields.com/white-noise-definition-vs-pink-noise/

Adriaensen, F. 2011. Loudness Measurement According to EBU-R128. In *Linux Audio Conference 2011. Maynooth, Ireland.*

Apple. 2012. Mastered for iTunes: Music as the Artist and Sound Engineer Intended. Technology Brief, Mastered for iTunes. Apple 2012. Accessed: February 2016, Quote p4.

Apple. 2016. Use Sound Check in iTunes to make song volume consistent. Turn on Sound Check in iTunes to make all of your play at the same volume. Quote p1.
Available at: https://support.apple.com/en-za/HT201724

Asikhule. 2013. New Audio Requirements for DSTV Ads. Accessed: October 2016. Available at: http://www.asikhule.com/resources/Asikhule-Update-Q3-2013.pdf

Asikhule. 2014. New R128 Requirements for Ads! 1st August 2013. Accessed: October 2016.
Available at: http://www.asikhule.com/about/blog/files/673baa26617348db1d5594dc29c88d6d-3.html

Auphonic. 2016. Auphonic Leveler Batch Processor. Accessed: August 2016.
Available at: https://auphonic.com/leveler

Bjorkman, M. & Rylander, R. 1997. Maximum Noise Levels in City Traffic. In *Journal of Sound and Vibration, 205 (4), 513-516.* Figure from p514.

Boley, J., L.M. & Danner, C. 2010. Measuring Dynamics: Comparing and Contrasting Algorithms for the Computation of Dynamic Range. In *AES Convention Paper 8178, Presented at the 129th Convention, November 2010 San Francisco.*

Cabot, R. & Dennis, I. 2011. Understanding and Verifying Loudness Meters. Technical Note, Technical Information from Qualis Audio. Figure from p2.

Camerer, F. 2010. On the way to Loudness nirvana: Audio levelling with EBU R128. EBU Technical review. 2010. Quote p1, Figure p3.

CarBlueprints. 2008. Car Blueprints, Renault Clio III, 2006 Model. Accessed: September 2016.
Available at: http://carblueprints.narod.ru/

Davis, D., P.E. & Brown, P. 2013. *Sound System Engineering*. U.K.: Focal Press. Chapter 7: Using the Decibel, p87-110.

DSP-Mobile. 2012. DSP Mobile Analyzer. Accessed: August 2016.
Available at: http://dspmobile.de/2012/10/analyzer-2-0/

DynamicRangeDatabase. 2014. Dynamic Range Database Album List.
Available at: http://dr.loudness-war.info/

EBU. 2010. Recommendation R128, Loudness Normalisation and Permitted Max Level of Audio Signals. EBU-UER, EBU Recommendation, Geneva 2010.

EBU. 2014. Recommendation R128, Loudness Normalisation and Permitted Max Level of Audio Signals. EBU-UER, EBU Recommendation, Geneva 2014.

EBU. 2016. Guidelines for production of programmes in accordance with EBU R128. Operating Eurovision and Euroradio. Version 3.0 Geneva. Quote p39.

FaberAcoustical. 2016. Sound Meter. Accessed: September 2016.
Available at: http://www.faberacoustical.com/apps/ios/soundmeter/

Fleischer, G. 2008. Strategies of the Hearing System Against Noise and Auditory Damage. Reprint from Reflections on sound. Norwegian University of Science and Technology. Figures from p93, 104, 113.

Fleischhacker, S. 2014. Audio Loudness Analysis. Technical white paper. Sencore. Figure from p6.

Fletcher, H. & Munson, W. 1933. Loudness, Its Definition, Measurement and Calculation. Bell Telephone Laboratories, Volume 5, August 1933.

Github. 2016. Audionuma, R128x a tool for loudness measurement of files on Mac OSX Intel. Accessed September 2016.
Available at: https://github.com/audionuma/r128x

Hadi, E. 2010. Loudness Control for Smart Environments. Massey University, School of Engineering and Advanced Technology. Quote p10, Figure p10.

Howard, D. & Angus, J. 2009. *Acoustics and Psychoacoustics*. Oxford: Focal Press. Chapter 2: Introduction to Hearing, p74-119, Figures p92, 93.

ITU-R. 2011. Recommendation ITU-R BS.1770-2: Algorithms to Measure Audio Programme Loudness and True-Peak Audio Level, BS Broadcasting Service (Sound). ITU-R Radio-communication Sector of ITU.

ITU-R. 2012. Recommendation ITU-R BS.1770-1: Requirements for Loudness and True-Peak Indicating Meters, BS Broadcasting Service (Sound). ITU-R Radio-communication Sector of ITU.

ITU-R. 2017. Recommendation ITU-R BS.1770-4: Algorithms to measure audio programme loudness and true-peak audio level, BS Broadcasting Service (Sound). ITU-R Radio-communication Sector of ITU.

Kardous, C. & Shaw, P. 2015. Do Sound Meter Apps Measure Noise Levels Accurately? Sound and Vibration, July 2015.

Levine, R. 2007. The Death of High Fidelity. Rolling Stone Magazine. Accessed: March 2016.
Available at: `http://www.rollingstone.com/news/story/17777619/the_death_of_high_fidelity`

Lewis, P. 1973. The Noise Generated by Single Vehicles in Freely Flowing Traffic. In *Journal of Sound and Vibration, 1973, 30 (2), 191-206.* Figure from p193.

Loots, J. 2016. An Empirical Investigation of Loudness Fluctuations in South African Broadcast Audio. University of Pretoria, Faculty of Humanities.

Lund, T. 2006. Control of Loudness in Digital TV. TC Electronic A/S, Risskov, Denmark, NAB BEC Proceedings. Figure: p57, Quotes: p57, p61.

Lund, T. 2009. Loudness Descriptors for Use in Production and Broadcast. VDT Symposium, Hohenkammer, TC Electronic A/S, Denmark. Figure p36.

Lund, T. 2013. Audio for Mobile TV, iPad and iPod. TC Electronic A/S, Risskov, Denmark.

Maempel, H. & Gawlik, F. 2009. The influence of sound processing on listeners' program choice in radio broadcasting. In *Audio Engineering Society. Convention Paper 7785. Presented at 126th Convention. May 2009, Munich, Germany.*

MeldaProduction. 2009. MLoudnessAnalyser. Accessed May, 2016.
Available at: `https://www.meldaproduction.com/MLoudnessAnalyzer`

MusicLoudnessAlliance. 2012. Loudness Normalisation: The future of file-based playback. Version 1.0 July 2012.

MusOpen. nd. Violin Concerto No. 3 in B minor, Op. 61. Accessed October 2016.
Available at: `https://musopen.org/sheetmusic/1548/camille-saint-saens/violin-concerto-no-3-in-b-minor-op-61/`

NCH. nd. Tone Generator Software. Accessed August 2016.
Available at: `http://www.nch.com.au/tonegen/`

Nielsen, S. & Lund, T. 2003. Overload in Signal Conversion. AES 23rd International Conference, Copenhagen, Denmark May 2003.

Nocross, S., L.M. & Thibault, L. 2011. A Status Report on Loudness Control Technologies and Standardisation. Advanced Audio Systems Communications Research Centre Canada, Submitted to the Canadian Radio-Television and Telecommunications Commission.

Nor, M., F.M.N.H. & Ariffin, A. 2008. Index for Vehicle Acoustical Comfort Inside a Passenger Car. In *Applied Acoustics 69 (2008) 343-353.*

NugenAudio. 2014. *MasterCheck Operation Manual.*

Nygren, P. 2009. Achieving Equal Loudness Between Audio Files - Evaluation and Improvements of Loudness Algorithms. Master's thesis, Royal Institute of Technology, Sweden.

Orban, B. 2000. Maintaining Audio Quality in the Radio Plant. Orban, A Harman International Company.

Orban, R. & Foti, F. 2001. What Happens to My Recording When it's Played on the Radio. In *Audio Engineering Society. Convention Paper 5469. Presented at 111th Convention. September 2001, New York, USA.*

Ormuz, K. & Muftic, O. 2004. Main Ambient Factors Influencing Passenger Vehicle Comfort. In *Proceedings of 2nd International Ergonomics Conference, Ergonomics, Croatia 2004.* Quote p77.

Ouis, D. 2001. Annoyance From Road Traffic Noise. In *Journal of Environmental Psychology, 21, 101-120.*

Plack, C. 2004. *Auditory Perception.* Psychology Press Ltd. Psychology: An International Perspective.

RadioShack. 2011. Radio Shack Digital Sound Level Meter, 33-2055, User's Guide. Accessed: September 2016.
Available at: `https://partydj.be/PDF-files/Handleidingen/Sound/Sphynx/sphynx_db-meter-manual.pdf`

Riedmiller, J., L.S. & Robinson, C. 2003. Intelligent Program Loudness Measurement and Control: What Satisfies Listeners? In *Dolby Laboratories Inc. USA. AES Convention Paper. Presented at 11th Convention October 2003, New York.* Quote p5, Figure p5.

Robjohns, H. 2014. The End of the Loudness War? Sound on Sound Magazine, Volume 29 Issue 4. P112-128.

Rogers, S. 2011. The Loudness War is Fought with and over Compression. In *Berklee College of Music. 131st AES New York, 2011.*

Skovenborg, E. & Lund, T. 2008. Loudness Descriptors to Characterise Programmes and Musical Tracks. In *AES Convention Paper, Presented at 125th Convention, October 2008 San Francisco.* Figure p3.

Skovenborg, E. & Lund, T. 2009. Loudness Descriptors to Characterise Wide Loudness-Range Material. In *AES Convention Paper, Presented at 127th Convention, October 2009 New York.* Quote p2.

Skovenborg, E. & Nielsen, S. 2004. Evaluation of Different Loudness Models with Music and Speech Material. In *AES Convention Paper 6618. Presented at 117th Convention, October 2004 San Francisco.*

Skovenborg, E. & Nielsen, S. 2007. Real-Time Visualisation of Loudness Along Different Time Scales. In *Digital Audio Effects International Conference, Bordeaux, France, 2007.*

SMPTE. 2016. Top SMPTE Standards. Society of Motion Pictures and Television Engineers, About SMPTE. Accessed: August 2016.
Available at: `https://www.smpte.org/top-standards`

Soulodre, G. & Lavoie, M. 2005. Stereo and Multichannel Loudness Perception and Metering. In *Communications Research Centre, Canada. AES Convention Paper 6618. Presented at 119th Convention, October 2005 New York.*

Southall, N. 2006. Imperfect Sound Forever. Accessed: March 2016.
Available at: `http://www.stylusmagazine.com/articles/weekly_article/imperfect-sound-forever.htm`

Speer, B. 2001. What Happened to Dynamic Range? Bob Speer's CD Mastering Services, CD Mastering Services, 2001-2016. Accessed: April 2016.
Available at: `http://www.cdmasteringservices.com/dynamicrange.htm`

Spikofski, G. & Klar, S. 2004. Levelling and Loudness in Radio and Television Broadcasting. EBU Technical Review, January 2004. Figure from p3.

Tagtaum. 2016. BeaTunes: ReplayGain. Accessed: February 2016.
Available at: `https://www.beatunes.com/en/itunes-replay-gain.html`

Thomann. 2016. MicWi436. Accessed: August 2016.
Available at: `https://www.thomann.de/gb/micw_i436.htm`

Urick, B. 2016. Pink Floyd: The Wall, A Complete Analysis. Accessed: October 2016.
Available at: `http://www.thewallanalysis.com/`

Vickers, E. 2010. The Loudness War: Background, Speculation and Recommendations. In *AES Convention Paper, Presented at 129th Convention, November 2010 San Francisco.* Figure p6.

Vickers, E. 2011. The Loudness War: Do Louder, Hypercompressed Recordings Sell Better?. In *Journal of the Audio Engineering Society, Vol 59, No 5, May 2011.*

WavesAudio. 2016. Trans-X. Accessed: September 2016.
Available at: `http://www.waves.com/plugins/transx#control-drum-transients-with-trans-x`

Wolters, M, M.H. & Riedmiller, J. 2010. Loudness Normalisation in the age of Portable Media Players. In *Dolby Germany and Dolby Laboratories Inc USA. AES Convention Paper, Presented at 128th Convention May 2010, London UK.*